

THE ANALYTICAL FOUNDATIONS
OF CELESTIAL MECHANICS

PRINCETON MATHEMATICAL SERIES

Editors: MARSTON MORSE, H. P. ROBERTSON, A. W. TUCKER

1. The Classical Groups
Their Invariants and Representations

BY HERMANN WEYL

2. Topological Groups

BY L. PONTRJAGIN

3. An Introduction to Differential Geometry
with Use of the Tensor Calculus

BY LUTHER PFAHLER EISENHART

4. Dimension Theory

BY WITOLD HUREWICZ AND HENRY WALLMAN

5. The Analytical Foundations of Celestial Mechanics

BY AUREL WINTNER

6. The Laplace Transform

BY DAVID VERNON WIDDER

7. Integration

BY EDWARD J. MC SHANE

8. Theory of Lie Groups

BY CLAUDE CHEVALLEY

9. Mathematical Methods of Statistics

BY HARALD CRAMÉR

THE ANALYTICAL FOUNDATIONS OF CELESTIAL MECHANICS

BY

AUREL WINTNER

1947

PRINCETON, NEW JERSEY
PRINCETON UNIVERSITY PRESS
LONDON: GEOFFREY CUMBERLEGE
OXFORD UNIVERSITY PRESS

IIA Lib.,

Copyright, 1941, by Princeton University Press

Second printing, 1947

Third printing, 1952

Printed in the United States of America

LEON LICHTENSTEIN
IN MEMORIAM

PREFACE

It was more than twelve years ago that, at the suggestion of the late Professor Lichtenstein, I began working on a book on the problem of three bodies. The original plan was to present a systematic account of the methods and results of the theory of the periodic and related particular solutions of the restricted problem of three bodies and its extensions, and to arrange everything else around these fundamental solutions.

However, during the progress of the work it became more and more clear that a systematic presentation of the mathematical theory of periodic solutions and their applications to the problem of the solar system, on the one hand, and to Strömgren's numerical investigations, on the other hand, must be preceded by a modernized treatment of those analytical aspects of the general theory of canonical systems which were originated by, and are still fundamental for, Celestial Mechanics as a whole. Through repeated discussions of the plan of the book with Professor G. D. Birkhoff, I became still more convinced of the necessity of such an approach. I am greatly indebted to him for the friendly and helpful interest which he has always taken in this book.

The title is intended to imply that the general topological methods in proofs of existence, as initiated by Poincaré, are not discussed in this volume. Nevertheless, this book could not have been written without the investigations of Levi-Civita and Birkhoff. Actually, the theory of periodic solutions will be illustrated only by the case of Hill's lunar theory; a case historically and methodically so fundamental as to necessitate an exception.

Approximately the first third of the book is based on a course of lectures on analytical mechanics, given for graduate students in physics and mathematics. It is therefore hoped that these chapters can serve as an introduction into the pure analysis of theoretical dynamics and of the theory of perturbations. Throughout the book (and especially in Chapter VI), I have tried not to repel that regrettable majority of younger mathematicians who have had no contact with theoretical astronomy.

Chapter I is perhaps unusual, in that it develops only the dynamical operators of canonical systems of differential equations, without disguising the actual content of the formalism by an introduction of

these equations themselves. In fact, the differential equations and their solutions are introduced only in Chapter II. Correspondingly, the theory of the canonical variation of constants in the theory of perturbations is not subordinated to the characteristic partial differential equation which, in fact, appears only as a by-product of the general theory of the transformations of phase space.

In Chapter II, emphasis is laid on a careful distinction between formal questions, which are always local in nature, and questions in the large, which are the actual problems of mathematical dynamics. While it is true that in most cases more is known about the possible nature of the non-local problems in Celestial Mechanics than about a workable approach to them, the sections dealing with the nature of non-local problems appeared to be rather necessary. In fact, without these sections it would have been hardly possible even to indicate in later chapters what, in case of n bodies, are actual problems and what must be considered at present pseudo-problems.

While Chapter I and Chapter II concern an arbitrary canonical system, Chapter III takes into account the peculiar quadratic structure of a dynamical Hamiltonian function. The only non-trivial case for which an explicit analytical formalism is available at present, namely, the case of two degrees of freedom, is considered in some detail in order that it may become available for application to the restricted problem of three bodies.

Chapter IV presents the problem of two bodies, as far as it is of theoretical interest and does not involve the practice of the determination of preliminary orbits. The treatment of this elementary case is focused on the fact that the Newtonian choice of the law of attraction is exceptional in every respect. While historical remarks are deferred to the end of the book in most cases, in this chapter it seemed to be advisable to put a few remarks of this nature into the text; for it is almost forgotten how much the theory of analytic functions, for instance, owes to the "elementary" problem of two bodies.

Chapter V is the longest chapter of the book. It is somewhat heterogeneous, since it attempts to give an account of our present knowledge of the problem of three or more bodies (with the exclusion of the theory of certain periodic and related motions). However, in a few cases I did not succeed in finding short-cuts to certain results for which lengthy proofs are available in original memoirs. In these few cases, I was content to mention (sometimes among the historical notes at the end of the book) the result without proof, but with an explanation of the rôle of the result or of the apparent reasons for

the difficulties of the proof; thus hoping to avoid a disruption of the scope of this volume by a reproduction of the lengthy original proof of usually isolated facts. On the other hand, I did not hesitate to point out problems which suggested themselves but to which I did not find a suitable approach. Actually, all that appears to be, at least in principle, in its definitive form at present is, on the one hand, Sundman's theory of binary and general collisions and, on the other hand, the theory of homographic solutions. Correspondingly, these two topics are treated in considerable detail.

Chapter VI, dealing with the restricted problem of three bodies, is relatively short only because the foundation for it has been sufficiently laid in the preceding chapters. The limitations to which this chapter had to be subjected were indicated at the beginning of this Preface. The sections on lunar theory deal with the fundamental mathematical questions of the theory of the Moon and stop at the border of the still unknown land of the "small divisors" in classical Celestial Mechanics.

In the references, which are collected in an appendix, an attempt was made to correct certain traditional injustices. In fact, even the classical literature of the great century of Celestial Mechanics appears to be saturated with rediscoveries (sometimes *bona fide* and sometimes not assuredly so); rediscoveries which, during the last hundred years, have somehow succeeded in establishing definite claims on discovery. The situation is often so involved that the subject deserves a detailed and precise historical study. Such a monographic completeness is not, of course, the task of the appendix, which indeed is likely to contain blunders (the more so as the literature before Lagrange was available to me only to a small extent).

In case of duplications in relatively modern literature, the references mention only the author whom I thought to be the first discoverer of the result or the method at hand. I had to decide on this procedure, after finding that, for instance, Pizzetti's theory of homographic solutions was repeatedly rediscovered within a quarter of a century; while Gascheau's result on the characteristic exponents of the equilateral solutions of relative equilibrium has accumulated at least five rediscoveries since his note appeared in the *Comptes Rendus* (1843). How inevitable such duplications are can be fully appreciated only by realizing the vastness of both the astronomical and mathematical literature of the problem of n bodies; a literature usually restricted to a very small public but spread from the beginning over many periodicals in many countries. In addition, in-

genious and perfectly satisfactory considerations often occur in publications which were not written by ϵ -trained mathematicians, and so a critical reader can easily become too irritated to look at the paper with care. Of course, the opposite situation is just as frequent.

I wish to thank Professor E. K. Haviland of Lincoln University, Drs. E. R. van Kampen and R. B. Kershner of Johns Hopkins University, and Dr. Wilfred Kaplan of the University of Michigan for their devoted reading either of the manuscript or of the proofs, and for their helpful suggestions and valuable advice.

I am indebted to the Committee on the Revolving Fund of the American Mathematical Society and to the National Research Council for making possible the publication of this volume.

Tamworth, N. H., September 1940

AUREL WINTNER

CONTENTS

page

CHAPTER I. DYNAMICAL OPERATIONS

§1–	§8.	Transformations.....	3
§9–	§14.	Lagrangian derivatives.....	9
§15–	§25.	The phase space.....	13
§26–	§38.	Canonical transformations.....	22
§39–	§46.	Canonical transformations and Pfaffians.....	28
§47–	§56.	Extended coordinate transformations.....	34
§57–	§64.	Canonical matrices.....	43
§65–	§78.	Rotations.....	49

CHAPTER II. LOCAL AND NON-LOCAL QUESTIONS

§79–	§90.	Local notions.....	58
§91–	§102.	Hamiltonian and Lagrangian systems.....	67
§103–	§118.	Solutions and canonical transformations.....	75
§119–	§130.	Non-local notions.....	85
§131–	§136.	Points of stability.....	98
§137–	§154.	Characteristic exponents.....	102

CHAPTER III. DYNAMICAL SYSTEMS

§155–	§166.	Hamiltonian and Lagrangian equations.....	112
§167–	§184.	Isoenergetic reduction.....	119
§185–	§193.	Single degree of freedom.....	131
§194–	§205.	Integrable systems.....	138
§206–	§226.	Systems with radial symmetry.....	147
§227–	§240.	Two degrees of freedom.....	161

CHAPTER IV. THE PROBLEM OF TWO BODIES

§241–	§257.	The solution paths.....	178
§258–	§273.	The anomalies.....	192
§274–	§284.	Expansions of the elliptic motion into Fourier series..	201
§285–	§299.	Expansions according to powers of the eccentricity..	212
§300–	§312.	Synodical coordinates.....	222

CHAPTER V. THE PROBLEM OF SEVERAL BODIES

§313–	§321.	Newton's law of gravitation.....	233
§322–	§332.	Consequences of the conservation integrals.....	242
§333–	§339.	Simultaneous collisions.....	251
§340–	§347.	Heliocentric coordinates.....	257
§348–	§354.	Binary collisions.....	266
§355–	§368.	Central configurations.....	273
§369–	§374.	Homographic solutions.....	284
§375–	§382.	Homographic solutions and central configurations...	295

§383-§389.	Elimination of the linear momentum.....	306
§390-§406.	Elimination of the angular momentum.....	315
§407-§414.	Real singularities.....	324
§415-§425.	The function-theoretical character of the collisions..	330
§426-§440.	The problem of three bodies.....	338

CHAPTER VI. INTRODUCTION TO THE RESTRICTED PROBLEM

§441-§445.	The restricted problem of three bodies.....	347
§446-§461.	Regularization.....	351
§462-§468.	The syzygical potential curve.....	359
§469-§477.	The potential surface.....	366
§478-§488.	The non-planar restricted problem.....	373
§489-§502.	Lunar systems.....	379
§503-§515.	Periodic lunar orbits.....	388
§516-§529.	Lunar theory.....	400

HISTORICAL NOTES AND REFERENCES

Chapter I	(§1- §78).....	413
Chapter II	(§79-§154).....	414
Chapter III	(§155-§240).....	416
Chapter IV	(§241-§312).....	421
Chapter V	(§313-§440).....	425
Chapter VI	(§441-§529).....	436

INDEX.....	445
------------	-----

THE ANALYTICAL FOUNDATIONS
OF CELESTIAL MECHANICS

CHAPTER I

DYNAMICAL OPERATIONS

Transformations.....	§ 1-§ 8
Lagrangian derivatives.....	§ 9-§14
The phase space.....	§15-§25
Canonical transformations.....	§26-§38
Canonical transformations and Pfaffians.....	§39-§46
Extended coordinate transformations.....	§47-§56
Canonical matrices.....	§57-§64
Rotations.....	§65-§78

Transformations

§1. An ordered collection of a finite number of scalars a_j will be called a vector $a = (a_j)$. By an m -vector will be meant a vector with m components. The latter will be thought of as arranged in the form of a "column," i.e., in the form

$$a = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix},$$

and not as a "row" (a_1, \dots, a_m) . If $b = (b_j)$ is another m -vector, $a \cdot b = b \cdot a$ will denote the scalar product $\sum a_j b_j$. If α is a scalar, the product $\alpha a = a \alpha$ denotes the m -vector (αa_j) .

By $C = (c_k^i)$ will be meant a matrix with an equal number of rows and columns. If this number is m , the matrix will be called an m -matrix. In $C = (c_k^i)$, the superscript i and the subscript k are thought of as the indices of the i -th row and the k -th column, respectively. If α is a scalar, the product $\alpha C = C \alpha$ denotes the m -matrix (αc_k^i) . Reserving the prime ' for the symbol of total differentiation with respect to a time variable t , the operation of transposition of $C = (c_k^i)$ will be denoted by a prime `; so that $C^{\text{'}} = (c_i^k)$. Thus, $C^{\text{'}} = C$ means that C is symmetric. And $B^{\text{'}} A^{\text{'}}$ is the transposed matrix of the product AB ($\neq BA$) of two m -matrices. The determinant of C will be denoted by $\det C$; so that $\det C \neq 0$ characterizes a non-singular C , i.e., a C for which the reciprocal matrix C^{-1} exists. The unit matrix will be denoted by $E = (e_k^i)$; so that $e_i^i = 1$, while $e_k^i = 0$ for $k \neq i$.

If A is an m -matrix and a an m -vector, Aa will denote the m -vector

into which a is sent by the linear transformation A . On the other hand, aA will be considered as undefined.

Needless to say, ABc will denote, for a pair A, B of m -matrices A, B and for an m -vector c , the m -vector Ca , where $C = AB$. Similarly, $a \cdot Cb$ will denote, for a pair a, b of m -vectors a, b and for an m -matrix C , the scalar $a \cdot c$, where $c = Cb$. By the definition of the transposed matrix, $a \cdot Cb = b \cdot C^*a$.

By 0 will be denoted not only the number zero but also the zero vector and the zero matrix.

§2. All numbers, variables and functions occurring will be understood to be real-valued, unless it is stated or implied that what is meant is the complex field.

A set \mathbf{D} in the Cartesian space of a variable m -vector $x = (x_i)$ will be called a domain if it is an open, connected, non-vacuous set.

A scalar, vector or matrix function $f = f(x)$ of x is called of class $C^{(\nu)}$, where ν is a fixed positive integer, if f is, on the domain \mathbf{D} under consideration, a (single-valued) function for which all partial derivatives of order not greater than ν exist and are continuous on \mathbf{D} . When no misunderstanding is possible, \mathbf{D} will not be mentioned explicitly. The class $C^{(\nu)}$ contains the class $C^{(\nu+1)}$.

If $f = f(x)$ is a scalar, vector or matrix function of class $C^{(1)}$, and x_i one of the components of the m -vector $x = (x_i)$, the scalar x_i , when written as a subscript of f , will denote partial differentiation with respect to x_i . On the other hand, the m -vector $x = (x_i)$ will be applied as a subscript of f only when the function $f(x)$ is either a scalar or an m -vector. In the first case, where f is a scalar function, the symbol $f_x \equiv f_x(x)$ will denote the gradient of f with respect to x ; so that f_x is the m -vector function (f_{x_j}) whose j -th component is the partial derivative f_{x_j} . In the second case, where $f = (f_k)$ is an m -vector function of the m -vector $x = (x_k)$, the symbol $f_x \equiv f_x(x)$ will be meant to be the m -matrix whose i -th row consists of the components of the gradient of the scalar function f_i with respect to $x = (x_k)$.

§3. It is clear that, for a given m -vector function $y = y(x)$ of class $C^{(1)}$, there exists a scalar function $s = s(x)$ whose gradient $s_x(x)$ is $y(x)$, if and only if $y(x)$ satisfies the integrability condition expressed by the symmetry, $y_x^i = y_x^i$, of the Jacobian matrix; y_x being the Hessian matrix, $(s_{x_i x_k}) = (s_{x_k x_i})$, of the scalar function $s = s(x)$ of class $C^{(2)}$. This only means that $y_x^i = y_x^i$ is the condition for the identical vanishing of the curl* of $y(x)$.

* By curl $y \equiv \text{curl } y(x)$ will be meant the m -matrix function $y_x - y_x^i$ of x ,

It follows that if a given m -matrix function $A = A(x)$ of class $C^{(1)}$ has the property that there exists for *every* scalar function $f = f(x)$ of a given class $C^{(\nu)}$ a scalar function $\bar{f} = \bar{f}(x)$ such that $\bar{f}_x = Af_x$ (i.e., if the matrix $A(x)$ transforms every gradient vector into a gradient vector), then $A(x) \equiv \mu E$, where μ is a scalar independent of $x = (x_i)$ and E is the unit matrix.*

§4. Suppose that, besides the m -matrix $A = A(x)$, there is given an m -vector function $a = a(x)$, and that the pair A, a has the property that one can find for *every* scalar function $f(x)$ of a given class $C^{(\nu)}$ a scalar function $\bar{f} = \bar{f}(x)$ such that $\bar{f}_x = a + Af_x$. Then, if $\bar{f} = g$ belongs to $f \equiv 0$, one sees that a is the gradient of g ; so that Af_x is, for every f , a gradient (namely, the gradient of $\bar{f} - g$). It follows, therefore, from §3 that $A(x) = \mu E$, where $\mu = \text{const.}$ Conversely, if there exist a scalar function $g(x)$ and a constant μ such that $a = g_x$ and $A(x) = \mu E$, then $\bar{f}_x = a + Af_x$ is satisfied by $\bar{f} = g + \mu f$ for every f .

Accordingly, the m -vector $a(x) + A(x)v(x)$ is, for a fixed pair a, A and for *every* gradient $v = f_x$, again a gradient, if and only if the given vector $a(x)$ is a gradient and the given matrix $A(x)$ is of the form μE , where E is the unit matrix and μ a scalar which does not depend on $x = (x_i)$.

§5. If $x = (x_i)$ and $y = (y_i)$ are two m -vectors, a mapping $y = y(x)$ of an x -domain on a y -domain will be called of class $C^{[\nu]}$, if the mapping is locally topological and such that both the function $y = y(x)$ and its locally unique inverse function $x = x(y)$ are of class

a matrix which is always skew-symmetric (and may, therefore, be replaced by an m -vector function of x only if $\frac{1}{2}m(m-1) = m$, i.e., in the usual case $m = 3$).

* In order to prove this, let $A_k(x)$ denote the m -vector representing the k -th column of $A(x)$. Since the vector $A(x)f_x(x)$ is required to be a gradient for every scalar polynomial $f = f(x) \equiv f(x_1, \dots, x_k, \dots, x_m)$, hence for every scalar polynomial $f = f(x_k)$ of the single variable x_k , it is seen, by placing $g(x_k) = f_{x_k}(x_k)$, that the vector $g(x_k)A_k(x)$, where k is arbitrary and $x = (x_i)$, is a gradient for every scalar polynomial $g(x_i)$ in the scalar x_i . It follows, therefore, from the integrability condition satisfied by vectors which are gradients, that each but the k -th component of m -vector $A_k(x)$ must vanish identically in $x = (x_i)$, and that the k -th component of $A_k(x)$ must be independent of every component of x except for the k -th. In other words, the m -matrix $A(x)$ must be a diagonal matrix in which the k -th diagonal element, say $\alpha_k = \alpha_k(x)$, is a function $\alpha_k(x_k)$ of the single component x_k of $x = (x_i)$. Consequently, the statement that $A(x) = \mu E$, where $\mu = \text{const.}$, is equivalent to the statement that $\alpha_i(x_i) = \alpha_k(x_k)$. Now, if the conditions $\alpha_i(x_i) = \alpha_k(x_k)$ were not satisfied, then the vector $A(x)f_x(x)$ could not be a gradient for the monomials $f(x) = x_i x_k$, where i, k are arbitrary.

$C^{(\nu)}$ in the sense of §2 (then the mapping $x = x(y)$ necessarily is of class $C^{[\nu]}$). A locally topological mapping $y = y(x)$ defined by a vector function $y(x)$ of class $C^{(\nu)}$ need not be of class $C^{[\nu]}$ (an example to this effect is, if $m = 1$, the mapping $y = x^3$ of $-1 < x < 1$ upon $-1 < y < 1$). By standard theorems concerning implicit functions, the mapping $y = y(x)$ is of class $C^{[\nu]}$ if and only if the function $y(x)$ is of class $C^{(\nu)}$ and has a non-vanishing Jacobian, $\det y_x(x)$, in the x -domain under consideration.

§6. Let $r = (r_i)$ be an n -vector and $H = H(p)$ a scalar function of class $C^{(2)}$, where $p = (p_i)$ is another n -vector. Suppose that the Hessian $\det (H_{p_i p_k}(p))$ does not vanish in the p -domain under consideration. Since this Hessian is the Jacobian of the gradient, $H_p(p)$, with respect to p , the mapping $r = r(p)$ defined by $r = H_p(p)$ is of class $C^{[1]}$. It turns out that the inverse mapping, $p = p(r)$, can be represented in the same form as the mapping $r = r(p) \equiv H_p(p)$, i.e., that there exists a scalar function $L = L(r)$ of class $C^{(2)}$ such that $p(r)$ is the gradient $L_r(r)$.

In order to prove this, define an $L = L(r)$ by placing

$$(1) \quad L(r) + H(p) = r \cdot p, \quad (r \cdot p = \sum r_i p_i; \text{ cf. §1}),$$

p being thought of as expressed in terms of the locally unique inverse, $p = p(r)$, of $r = r(p)$. Thus, $L(r) = r \cdot p(r) - H(p(r))$. Hence, $L_r(r) = p(r)$ plus two terms whose sum is 0, since $r - H_p(p(r)) = 0$ is, in virtue of $r = r(p) \equiv H_p(p)$, an identity in r . This proves that the locally unique inverse $p = p(r)$ of the mapping $r = H_p(p)$ can be represented by means of the scalar function $L(r)$ in the form $p = L_r(r)$. Since $r = H_p(p)$ is of class $C^{[1]}$, so is the inverse mapping $p = L_r(r)$; so that the function $L(r)$ is of class $C^{(2)}$. Finally, the product of the Hessian matrices of the scalar functions $L(r)$, $H(p)$ is

$$(2) \quad (L_{r_i r_k}(r))(H_{p_i p_k}(p)) \equiv E (= \text{unit matrix})$$

in virtue of either of the transformation formulae

$$(3) \quad r = H_p(p), \quad p = L_r(r).$$

In fact, these transformation formulae are inverses of each other and have, therefore, reciprocal Jacobian matrices. A consequence of (2) is that not only $\det (H_{p_i p_k}(p)) \neq 0$ in the p -domain but also $\det (L_{r_i r_k}(r)) \neq 0$ in the r -domain.

In the above proof for the existence of an L , the function $L(r)$ has been defined by means of (1). Actually, the requirement that the

locally unique inverse mapping $p = p(r)$ of $r = H_p(p)$ is $p = L_r(r)$, determines not the scalar function $L(r)$ itself but merely its gradient, $L_r(r)$; thus leaving undetermined an arbitrary additive constant in $L(r)$. This means that, while (1) is not an identity in p in virtue of $r = r(p) \equiv H_p(p)$, the difference between the two members of the equation (1) is a constant in virtue of either of the (equivalent) transformation formulae (3). In what follows, it will be assumed that this arbitrary additive scalar constant is chosen to be 0.

§7. Let Π denote the operation of permutation which replaces, in each of the assumptions and assertions of §6, the letters p, H and r, L by the letters r, L and p, H , respectively. Thus, Π replaces the assumption that there is given a function $H(p)$ of class $C^{(2)}$ with a non-vanishing Hessian, by the assertion that there exists a function $L(r)$ of class $C^{(2)}$ with a non-vanishing Hessian. Similarly, Π interchanges the assumption $r = r(p) \equiv H_p(p)$ and the assertion $p = p(r) \equiv L_r(r)$. Finally, (1), (2), (3) go over into themselves on the permutation Π . It follows that, instead of starting with an $H(p)$ and assigning the mapping $r = H_p(p)$, one can start with an $L(r)$ and assign the mapping $p = L_r(r)$. Then $H(p)$ is to be defined by means of (1), where r is thought of as expressed by means of the locally unique inverse, $r = r(p)$, of the given mapping $p = L_r(r)$ of class $C^{(1)}$ as a function of p .

Since two-fold application of Π clearly gives the identical permutation, and since the transformation formulae (3) are locally unique inverses of each other, it is seen that the correspondence between $H(p)$ and $L(r)$ is involutory. In other words, if $L(r)$ belongs to $H(p)$, then $H(p)$ belongs to $L(r)$. It follows that if $L_I(r)$ belongs to $H_I(p)$, and $H_{II}(p)$ to $L_I(r)$, finally $L_{II}(r)$ to $H_{II}(p)$, then $H_{II}(p) \equiv H_I(p)$ and $L_{II}(r) \equiv L_I(r)$.

§8. Suppose that one, hence both, of the two scalar functions H, L (of class $C^{(2)}$ and of non-vanishing Hessian in the respective n -vectors p, r) contains some l -vector s as a parameter, and that one, hence both, of the n -vector functions $H_p(p, s), L_r(r, s)$ of $n + l$ scalar variables is of class $C^{(1)}$ in these $n + l$ variables together. Thus, the formulae of §6 become

$$(4) \quad p = L_r(r, s), \quad r = H_p(p, s); \quad (5) \quad L(r, s) + H(p, s) = r \cdot p;$$

$$(6) \quad (L_{r_i r_k}(r, s))(H_{p_i p_k}(p, s)) \equiv E.$$

While $r = H_p(p, s)$ is, for every fixed point s of the l -dimensional

parameter domain, a mapping of class $C^{[1]}$ of the n -dimensional p -domain on the n -dimensional r -domain, the pair of relations, $r = H_p(p, s)$, $s = s$, defines a mapping of class $C^{[1]}$ of the $(n + l)$ -dimensional (p, s) -domain on the $(n + l)$ -dimensional (r, s) -domain. And the two transformation formulae (4) are reciprocal.

Thus, $r = r(p, s)$ and $p = p(r, s)$. Hence, on eliminating from (5) either r or p , say p , one sees that the variable n -vector r and the parameter l -vector s are connected by the scalar identity $L(r, s) + H(p(r, s), s) - r \cdot p(r, s) = 0$ in (r, s) . Differentiating this identity partially with respect to the components of the l -vector $s = (s_i)$, and then using the fact that $r - H_p(p, s)$ vanishes by (4), one obtains

$$(7) \quad L_s(r, s) + H_s(p, s) = 0,$$

the canceling being the same as in the calculation which, in §6, led from $L + H = p \cdot r$ to $L_r(r) = p(r)$.

It is understood that the gradient relation (7) is thus proved as an identity which holds in virtue of (4) and (5) together. It turns out, however, that (7) is an identity

- (i) in virtue of (4) alone and
- (ii) in virtue of (5) alone.

Since (7) is an identity in virtue of (4) and (5) together, and since (5) is, by the end of §6, an identity in virtue of (4) up to an additive constant, a constant which is removed by the gradient process leading from (5) to (7), it is clear that (7) is an identity in virtue of (4) alone. This proves (i). As to (ii), it is sufficient to observe that if the three vectors r, p, s are thought of as independent of each other, then the gradient of $r \cdot p$ with respect to s is 0; so that (7) is an identity in virtue of (5) alone.

It is similarly seen that the relations (4) hold

- (i) not only in virtue of (4), i.e., as relations defining the mapping of (p, s) on (r, s) , but also
- (ii) as identities in virtue of (5), i.e., as relations between the three vectors r, p, s which are subject to the single relation (5) only.

This ambivalence, (i)–(ii), in the possible interpretation of the gradient relations (7) and (4) is a fundamental property of the involutory transformation formulae (4), and is usually described by saying that the mapping (4) of (p, s) on (r, s) is a contact transformation. The word “contact” refers to partial differentiations of the *first* order. Notice that the relation (6) between the partial derivatives of the *second* order clearly is not an identity in virtue of (5) alone.

Lagrangian Derivatives

§9. Let \mathbf{R} , \mathbf{Q} be domains in the spaces of two n -vectors $r = (r_i)$, $q = (q_i)$, respectively, and let \mathbf{T} be a domain of a scalar variable t . Let $L = L(r, q; t)$ be a scalar function such that the n -vector function $L_r(r, q; t)$ is of class $C^{(1)}$ on the product space* of \mathbf{R} , \mathbf{Q} , \mathbf{T} . Denote by a prime ' total differentiation with respect to the "time" t .

Let $q(t)$ be an n -vector function of class $C^{(2)}$ on \mathbf{T} such that the "path" $q = q(t)$ in the q -space is situated within \mathbf{Q} , and the "velocity" $r = q'(t)$ within \mathbf{R} , for all t contained in the t -interval \mathbf{T} . Then one can define on \mathbf{T} a continuous n -vector function $[L]_q = ([L]_{q_i})$ of t by placing

$$(1) \quad [L]_q = L'_{q'} - L_q, \text{ i.e., } [L]_{q_i} = L'_{q'_i} - L_{q_i}, \quad (i = 1, \dots, n),$$

where $L(q', q; t)$ is thought of as expressed as a function of t . The subscripts q , q_i of the symbols $[]_q$, $[]_{q_i}$ do not denote partial differentiations but belong to these symbols. Thus, the i -th component of the n -vector $[L]_q$ is

$$(2) \quad [L]_{q_i} = \sum q'_k L'_{q'_i q'_k} + \sum q'_k L'_{q'_i q_k} + L'_{q'_i t} - L_{q_i},$$

where both summations \sum run from $k = 1$ to $k = n$, and the subscripts q'_i , q_i , t on the right of (2) denote partial differentiations of

$$(3) \quad L = L(q', q; t), \quad \text{where } q' = (q'_i), \quad q = (q_i).$$

The n -vector $[L]_q$ and its components $[L]_{q_i}$ will be referred to as the Lagrangian derivatives of L (along the parametrized path $q = q(t)$ in the q -space).

One easily verifies from (2) or (1) the scalar identity

$$(4) \quad (-L + q' \cdot L_{q'})' = -L_t + q' \cdot [L]_q, \quad (' = d/dt; a \cdot b = \sum a_i b_i).$$

§9 bis. The identity (4) suggests a hidden parallelism between t and the n -vector $q = (q_i)$. Introduce, therefore, an $(n+1)$ -vector $q = (q_k)$ by placing $q_0 = t$, $q_1 = q_1$, \dots , $q_n = q_n$, and put $L(q', q) \equiv L(q', q; t)$. Since $q_0 = t$, $q'_0 = 1$, $q''_0 = 0$, application of (2) to L shows that

$$(1 \text{ bis}) \quad [L]_{q_0} = -L_t; \quad [L]_{q_i} = [L]_{q_i}, \quad i = 1, \dots, n; \quad (L = L).$$

Hence, (4) appears in the symmetric form

$$(4 \text{ bis}) \quad (-L + q' \cdot L_{q'})' = q' \cdot [L]_q.$$

* By the product space of \mathbf{R} , \mathbf{Q} , \mathbf{T} is meant the set of those points of the $(2n+1)$ -dimensional $(r, q; t)$ -space for which r , q , t are points of \mathbf{R} , \mathbf{Q} , \mathbf{T} , respectively.

§10. Let $\bar{q} = \bar{q}(q; t)$ be an n -vector function which is of class $C^{(2)}$ in the $(n + 1)$ -dimensional $(q; t)$ -domain and has there a non-vanishing Jacobian with respect to q ; so that

$$(5) \quad q = q(\bar{q}; t)$$

is a mapping of class $C^{[2]}$ for every fixed t . The path $q = q(t)$ considered in §9 is mapped on a path $\bar{q} = \bar{q}(t)$ in such a way that, if Q denotes the Jacobian matrix of q with respect to \bar{q} at a fixed t , then

$$(6) \quad \begin{aligned} q' &= Q\bar{q}' + q_t, \text{ where} \\ Q &= q_{\bar{q}}, \det Q \neq 0, \text{ is an identity in } t \text{ in virtue of (5).} \end{aligned}$$

Starting with the L of §9, define an \bar{L} by the requirement that

$$(7) \quad \bar{L}(\bar{q}', \bar{q}; t) \equiv L(q', q; t) \text{ in virtue of (5) and (6).}$$

Then the differentiability assumptions made in §9 with regard to L are satisfied by \bar{L} . Furthermore, one can verify from the definitions (1) and (7) by straightforward differentiations that*

$$(8) \quad w = Q^{-1}\bar{w}, \text{ where } w = [L]_q, \bar{w} = [\bar{L}]_{\bar{q}}; Q = q_{\bar{q}} \text{ (} \det Q \neq 0 \text{)}.$$

§11. Needless to say, the function \bar{L} defined by (7) is not the function $L(\bar{q}', \bar{q}; t)$. Not even so much is true that if the transformation (5)–(6) is very close to the identical transformation $q = \bar{q}$, $q' = \bar{q}'$, then $L(\bar{q}', \bar{q}; t)$ is very close to $\bar{L}(\bar{q}', \bar{q}; t)$, i.e., to $L(q', q; t)$.

For let a transformation (5)–(6) be very close to $\bar{q} = q$, $\bar{q}' = q'$, in the sense that (5)–(6) can be embedded into a family†

$$(9) \quad \bar{q} = q + \epsilon f + o(\epsilon), \quad \bar{q}' = q' + \epsilon f' + o(\epsilon)$$

of such transformations, where $\epsilon > 0$ is a small parameter independent of q' , q , t , while $f = f(q', q; t)$ is a fixed n -vector function not con-

* The transformation rule (8) of the Lagrangian derivatives can be expressed by saying that, in virtue of (6) and (7), the n -vector $[L]_q$ behaves under a mapping (5) as if it were a covariant tensor in the q -space alone, and t did not occur in the transformation formula (5) of the coordinate vector q . On the other hand, the transformation rule (6) of the velocity vector is not that of a contravariant tensor in the q -space unless t does not occur in (5), i.e., unless $q_t \equiv 0$.

† By $o(\epsilon)$ is meant any function of q' , q , t , ϵ which has the property that the ratio $o(\epsilon):\epsilon$ tends, as $\epsilon \rightarrow 0$, uniformly to 0 for all q' , q , t under consideration.

Notice that, even if the function f is analytic, the first of the relations (9) does not imply the second, since the t -derivative of the first $o(\epsilon)$ need not be an $o(\epsilon)$; cf. the notion of a “weak neighborhood” in the calculus of variations

taining ϵ ; so that f is the partial derivative \bar{q}_ϵ at $\epsilon = 0$. To say that $L(\bar{q}', \bar{q}; t)$ is very close to $\bar{L}(\bar{q}', \bar{q}; t)$ is to say that

$$(10) \quad L(\bar{q}', \bar{q}; t) = \bar{L}(\bar{q}', \bar{q}; t) + o(\epsilon) \text{ in virtue of (9).}$$

And (10) is not true for every transformation (9) but only for transformations (9) in which the n -vector $f = f(q', q; t)$ satisfies, with reference to the given scalar function $L = L(q', q; t)$, a certain condition; namely, the condition that

$$(11) \quad (f \cdot L_{q'})' = f \cdot [L]_q, \text{ where } f = f(q', q; t), \quad L = L(q', q; t),$$

be an identity in the $(2n + 1)$ -dimensional $(q', q; t)$ -domain.

In fact, on substituting (9) into $L(\bar{q}', \bar{q}; t)$ and then denoting by $\lambda = \lambda(q', q; t)$ the partial derivative $L_\epsilon(\bar{q}', \bar{q}; t)$ at $\epsilon = 0$, one sees that $\lambda = f \cdot L_q + f' \cdot L_{q'}$. Hence, $\lambda = -f \cdot [L]_q + (f \cdot L_{q'})'$, by the definition (1). Thus, by Taylor's formula,

$$(12) \quad L(\bar{q}', \bar{q}; t) = L(q', q; t) + \epsilon \{ -f \cdot [L]_q + (f \cdot L_{q'})' \} + o(\epsilon),$$

since $L(\bar{q}', \bar{q}; t)$ and its derivative $L_\epsilon(\bar{q}', \bar{q}; t)$ reduce at $\epsilon = 0$ to $L(q', q; t)$ and $\{ \} = \lambda$, respectively. Now, while (12) holds for any $f = f(q', q; t)$ in (9), one sees from the definition (7), that the assumption (10) requires the vanishing of the coefficient $\{ \} = \lambda \equiv \lambda(q', q; t)$ of ϵ in (12). This proves that (11) is the condition imposed on f by the assumption (10).

§11 bis. It follows that if a family (9) of transformations is such as to leave $L(q', q; t)$ invariant for every ϵ , i.e., such that

$$(13) \quad L(\bar{q}', \bar{q}; t) \equiv L(q', q; t) \text{ in virtue of (9),}$$

then the n -vector $f(q', q; t) = (\bar{q}_\epsilon)_{\epsilon=0}$ satisfies the identity (11). In fact, (13) is sufficient for (10).

§12. A classical application of the consequence (11) of (13) will be mentioned in §96.

As another application, suppose that (13), or at least (10), is satisfied, and that

$$(14) \quad L_t \equiv 0, \quad f_t \equiv 0, \quad \text{i.e.,} \quad L = L(q', q), \quad f = f(q', q).$$

Let the path $q = q(t)$ considered in §9 be a closed curve in the q -space, i.e., $q(t) \equiv q(t + \tau)$ for some period $\tau > 0$. Then

$$(15) \quad 0 = \int_0^\tau f \cdot [L]_q dt,$$

the integrand being expressed as a function of t along the arbitrary closed path. In fact, the functions $f(q'(t), q(t))$, $[L(q'(t), q(t))]_q$ of t have the period τ ; so that (15) is clear from (11).

§13. If $L_t \equiv 0$, i.e., $L = L(q', q)$, then

$$(16) \quad 0 = \int_0^\tau q' \cdot [L]_q dt$$

for any path of the type considered in §12. In fact, (16) follows from (4) in the same way as (15) did from (11).

Incidentally, if the family (9) is defined by $\bar{q} = q(t + \epsilon)$, $\bar{q}' = q'(t + \epsilon)$, then $f \equiv (\bar{q}_\epsilon)_{\epsilon=0}$ becomes $q' = q'(t)$, and so (15) reduces to (16). This agrees with §9 bis.

§13 bis. If $L_t \equiv 0$ and if one joins two points $q = q^I$, $q = q^{II}$ of the q -domain by an oriented path $q = q(t)$ of class $C^{(2)}$, then the line integral

$$(17) \quad \int_{q^I}^{q^{II}} [L]_q \cdot dq$$

does not depend on the path but only on its end points q^I , q^{II} , and on the values of q' , q'' at these end points, provided that the paths considered are within a simply-connected domain. This is a mere restatement of (16).

§14. Instead of considering, as since §9, a single path $q = q(t)$, let $q = q(c; t)$ be a family of such paths, depending on a certain number, $m (\geq 1)$, of parameters c_j in such a way that the n -vector function $q = q(c; t)$ of the m -vector $c = (c_j)$ and of the time t is of class $C^{(2)}$ in the $(m + 1)$ -dimensional $(c; t)$ -domain under consideration. Let, in addition, $t^I = t^I(c)$ and $t^{II} = t^{II}(c)$ be two arbitrarily preassigned functions of class $C^{(1)}$ such that $(c; t^I)$ and $(c; t^{II})$ are in the $(c; t)$ -domain when c is in the c -domain. Then, $L = L(q', q; t)$ being the given scalar function considered since §9, the scalar

$$(18) \quad S = \int_{t^I}^{t^{II}} L dt \equiv \int_{t^I(c)}^{t^{II}(c)} L(q'(c; t), q(c; t); t) dt$$

is a function $S = S(c)$ of class $C^{(1)}$ in the c -domain.

According to the fundamental formula of the calculus of variations, one has for this function $S(c)$ of the m variables c_j the identity

$$\begin{aligned}
 (19) \quad \delta S = & \sum_{\nu=I}^{II} (-1)^\nu (L - q' \cdot L_{q'})_{t=t^\nu} \delta t^\nu \\
 & + \sum_{\nu=I}^{II} (-1)^\nu (L_{q'})_{t=t^\nu} \cdot \delta(q)_{t=t^\nu} - \int_{t^I}^{t^{II}} ([L]_q \cdot \delta q)_{t=t} dt.
 \end{aligned}$$

The operator δ occurring in (19) is defined by

$$(20) \quad \delta = \sum_{j=1}^m dc_j \frac{\partial}{\partial c_j}; \quad \text{so that} \quad dF(t; c) = F_t(c; t)dt + \delta F(c; t).$$

Thus, interchanging of the order of differentiations shows that $(\delta F)' = \delta(F')$. It is also seen from (20) that $\delta F = dF$, if F is a function of c alone; hence, $\delta S = dS$, $\delta t^I = dt^I$, $\delta t^{II} = dt^{II}$ in (19). On the other hand, $\delta t \neq dt$, since δF vanishes, by (20), for a function F of t alone, and so, in particular, for $F = t$.

Clearly, the sum of the five expressions in the representation (19) of $\delta S \equiv dS$ is a Pfaffian of the form $g_1 dc_1 + \cdots + g_m dc_m$, where $g_j = g_j(c_1, \cdots, c_m) \equiv g_j(c)$. Thus, (19) states merely that the coefficient $g_j(c)$ of this Pfaffian is identical with the partial derivative $S_{c_j}(c)$ of the function (18) of c (so that, in particular, the Pfaffian is a complete differential). Correspondingly, Lagrange's classical proof of (19) consists of a straightforward differentiation of (18) with respect to the c_j , followed by a partial integration which is based on the definition (1) of $[L]_q$ and on an application of $\delta(F') = (\delta F)'$ to $F(c; t) = q(c; t)$; cf. the proof of (12).

The Phase Space

§15. The assumption made, since §9, with regard to the scalar function $L(q', q; t)$ of the two n -vectors $r(= q')$, q and of the time t was that the n -vector function $L_r(r, q; t)$ be of class $C^{(1)}$ in the $(r, q; t)$ -domain under consideration. In what follows, it will, in addition, be assumed the Jacobian of the components L_{r_i} of L_r with respect to the components r_i of r , i.e., the n -rowed Hessian $\det(L_{r_i r_k})$ of L , vanishes at no point of the $(2n+1)$ -dimensional $(r, q; t)$ -domain. Then one can identify $L(r, q; t)$ with the function $L(r; s)$ considered in §8; the parameter vector s with an arbitrary number, l , of components s_1, \cdots, s_l being represented by the n components q_1, \cdots, q_n of q and by t ; so that $l = n+1$. Thus, on writing $q' = (q'_i)$ instead of $r = (r_i)$, one sees that the relations (4), (5), (6) of §8 go over into

$$(1_1) \quad p = L_{q'}(q', q; t);$$

$$(1_2) \quad q' = H_p(p, q; t);$$

$$(2_1) \quad L(q', q; t) + H(p, q; t) = q' \cdot p;$$

$$(2_2) \quad (L_{q'_i q'_k})(H_{p_i p_k}) = E,$$

while (7), §8 splits into the pair

$$(3_1) \quad L_q(q', q; t) + H_q(p, q; t) = 0;$$

$$(3_2) \quad L_t(q', q; t) + H_t(p, q; t) = 0.$$

The n -vector relations (1₁)–(1₂), (3₁) and the scalar relation (3₂) admit of the two-fold interpretation explained at the end of §8. According to (2₂), the two n -rowed determinant conditions

$$(4_1) \quad \det (L_{q'_i q'_k}(q', q; t)) \neq 0; \quad (4_2) \quad \det (H_{p_i p_k}(p, q; t)) \neq 0$$

are equivalent in virtue of the transformation formulae (1₁)–(1₂). The latter are, by §7, reciprocal and involutory.

The effect of the transformation (1₁) or (1₂) on the path $q = q(t)$ considered in §9 is that the path on $q = q(t)$ of class $C^{(2)}$ in the n -dimensional q -space and the velocity vector $q' = q'(t)$ along this path become replaced by a path $x = x(t)$ of class $C^{(1)}$ in a $2n$ -dimensional x -space, the latter space being formed by the $n + n$ components of two n -vectors $p = (p_i)$, $q = (q_i)$. In other words, $x = (x_i)$ is defined as the $2n$ -vector

$$(5) \quad x = (x_i): x_i = p_i, x_{i+n} = q_i; \quad \text{so that} \quad H(p, q; t) = H(x; t).$$

§16. The components p_i of the n -vector (1₁) are usually referred to as the “momenta” which are, with reference to the given function $L(r, q; t)$, “canonically conjugate” to the components $r_i = q'_i$ of the “velocity” vector $r = q'$. The n -dimensional q -space of the “coordinates” q_i is called the “configuration space,” the $2n$ -dimensional x -space defined by (5) the “phase space.” The integer n is the “degree of freedom.” Finally, L and H are called an associated pair of Lagrangian and Hamiltonian functions, respectively.

As far as the representation of the Lagrangian derivatives in Hamiltonian terms is concerned, it is clear from (1₁), (3₁), §15 and (1), §9 that

$$(6) \quad p' + H_q(p, q; t) = [L]_q; \quad \text{while} \quad -q' + H_p(p, q; t) = 0,$$

by (1₂). Similarly, (1₁), (2₁), (3₂) show that (4), §9 is equivalent to $H' - H_t = q' \cdot [L]_q$.

§16 bis. Instead of the Lagrangian function $L(q', q; t)$ with n de-

degrees of freedom, consider, for a moment, the Lagrangian function L^* defined by means of (5) as

$$(7) \quad L^*(x', x; t) = -H(x_1, \dots, x_{2n}; t) + \sum_{i=1}^n x_i x'_{i+n}.$$

Thus, L^* has $2n$ degrees of freedom but contains only n of the $2n$ velocity components x'_j , where the $2n$ components x_j of the vector x of the phase space are considered as forming the components of a $2n$ -dimensional configuration space.

Application of the definition $[L^*]_{x_j} = L^*_{x'_j} - L^*_{x_j}$, ($j = 1, \dots, 2n$), of a Lagrangian derivative to (7) gives

$$(8) \quad [L^*]_{x_i} = H_{x_i}(x; t) - x'_{i+n}, \quad [L^*]_{x_{i+n}} = H_{x_{i+n}}(x; t) + x'_i, \\ (i = 1, \dots, n).$$

Comparing (8) with (5), one sees that the pair of n -vector identities (6) can be written in the rather symmetric form

$$(9) \quad -q'_i + H_{p_i} = [L^*]_{p_i}, \quad p'_i + H_{q_i} = [L^*]_{q_i}, \quad (i = 1, \dots, n).$$

Furthermore, comparison of (7) with (2₁) shows that $L^* = L$ in virtue of (5).

§17. It is clear from (6), §10 and (1₁), §15 that if the configuration space is, for varying t , subject to a point transformation (5), §10, the corresponding point transformation of the phase space is uniquely determined. Such extensions of transformations of the n -dimensional q -space to transformations of the $2n$ -dimensional x -space will be studied in §48.

In what follows, a more general case will be considered, namely, the case of transformations of the x -space which need not be derivable from transformations of the q -space. Thus, if y denotes the $2n$ -vector into which x is transformed, the transformations to be considered are of the type $y = y(x; t)$, where, corresponding to (5),

$$(10) \quad y = (y_j): \quad y_i = u_i, \quad y_{i+n} = v_i, \\ (i = 1, \dots, n; j = 1, \dots, 2n);$$

$u = (u_i)$ and $v = (v_i)$ denoting the n -vectors which represent the new momenta and coordinates, respectively.

It will be assumed that the $2n$ -vector function $y(x; t)$ has a Jacobian $2n$ -matrix $y_x(x; t)$ which is of class $C^{(1)}$ and of non-vanishing

determinant in the $(2n + 1)$ -dimensional $(x; t)$ -domain under consideration; so that the mapping

$$(11_1) \quad y = y(x; t); \quad (11_2) \quad x = x(y; t)$$

of the two phase spaces x, y on each other is of class $C^{[1]}$ for every fixed t . In virtue of these transformation formulae, a function of the position in the $(x; t)$ -space becomes a function of the position in the $(y; t)$ -space, and conversely. For instance, if $F = F(y; t)$ is a scalar function of class $C^{(1)}$, partial differentiation shows that

$$(12) \quad F_x = \Gamma' F_y, \quad \text{where} \quad \Gamma = y_x, \quad (\det \Gamma \neq 0),$$

is an identity in virtue of (11_1) or (11_2) . Caution is necessary only in case of a partial derivative F_t with respect to the time.* In fact, $F_t(y; t)$, where y is fixed, is in virtue of (11_1) not the same thing as $F_t(y(x; t); t)$, where x is fixed.

The Jacobian $2n$ -matrix Γ occurring in (12) will be thought of as expressed by means of (11_2) as a function $\Gamma(y; t)$ of $(y; t)$, unless the contrary is said; and Γ_t will denote the matrix obtained by partial differentiation of the $4n^2$ elements of $\Gamma(y; t)$ with respect to t for a fixed y . On the other hand, by y_t will be meant the $2n$ -vector $y_t(y; t)$ which results if one differentiates $y(x; t)$ partially with respect to t for a fixed x and then expresses x by means of (11_2) in terms of $(y; t)$. Thus,

$$(13_1) \quad y_t = y_t(y; t); \quad (13_2) \quad y_x = \Gamma = \Gamma(y; t), \quad (\det \Gamma \neq 0).$$

Assuming that y_t also is of class $C^{(1)}$, one finds by straightforward differentiations (cf. §2, §1) that, in virtue of the transformation formulae (11_1) – (11_2) of class $C^{[1]}$,

$$(14_1) \quad (y_x)_t = (y_t)_y (y_x); \quad (14_2) \quad y_x = (x_y)^{-1}; \quad (14_3) \quad y' = (y_x)x' + y_t;$$

(14_1) , (14_2) being identities in the phase space for every fixed t , while (14_3) is an identity in t along any path $y = y(t)$ or $x = x(t)$ of class $C^{(1)}$ in the phase space. Using (13_1) , (13_2) , one can write (14_1) , (14_2) , (14_3) as

$$(15_1) \quad (y_t)_y = \Gamma_t \Gamma^{-1} \quad (15_2) \quad x_y = \Gamma^{-1} \quad (15_3) \quad y' = \Gamma x' + y_t.$$

§18. Needless to say, all these identities hold also when x or y are vectors not in a $2n$ -dimensional phase space but in a space of arbi-

* Cf. the “Lagrangian” and “Eulerian” points of view in the kinematics of continua.

trary dimension number, m . In this sense, (11₁) and (15₃) are not different from (5), §10 and (6), §10; while (12), (15₁), (15₂) might have been used in the verification of the identity (8), §10.

If F^1, \dots, F^l are l scalar functions of class $C^{(1)}$ which depend on the m -vector x and possibly on t , let

$$(15 \text{ bis}) \quad (F_x^1, \dots, F_x^l)$$

denote the "Jacobian matrix" in which the columns are the gradients of the F with respect to x ; so that (15 bis) has m rows and l columns. The l functions F are said to be independent in the domain under consideration if (15 bis) is of rank l in this domain,* i.e., if there exists for every point of the domain a non-vanishing minor with l rows and l columns (this implies that $l \leq m$).

A function will be called of the conservative type if it does not contain the time explicitly. For instance, a transformation (11₁)–(11₂) is called conservative if $y = y(x)$, hence $x = x(y)$. Accordingly, conservative functions of x are sent by conservative transformations into conservative functions of y . It is seen from (3₂), §15, that if the Lagrangian function is of the conservative type $L = L(q', q)$, then so is the Hamiltonian function $H = H(p, q)$, and conversely.

§19. Suppose that $m = 2n$, and denote by (e_k^i) and (0) the unit and zero n -matrices, respectively. Let I denote the constant $2n$ -matrix†

$$(16) \quad I = \begin{pmatrix} (0) & (e_k^i) \\ - (e_k^i) & (0) \end{pmatrix}; \text{ so that } I' = -I, I^{-1} = -I, \det I = 1.$$

By (16), (5) and (6), the $2n$ -vector relation

$$(17) \quad x' + IH_x \equiv \begin{pmatrix} p' \\ q' \end{pmatrix} + I \begin{pmatrix} H_p \\ H_q \end{pmatrix} \equiv \begin{pmatrix} p' + H_q \\ q' - H_p \end{pmatrix} = \begin{pmatrix} [L]_q \\ 0 \end{pmatrix},$$

where $p, q, H_p, H_q, [L]_q$ and 0 are n -vectors, is an identity in t .

* According to a theorem, now standard, the above definition of independence coincides with the classical notion of independence if one disregards nowhere dense sets in the x -space.

† This skew-symmetric matrix, which will play a fundamental rôle in what follows, is known to represent the normal form of an arbitrary non-singular skew-symmetric bilinear form; in the sense that there exists for every non-singular skew-symmetric matrix S a non-singular matrix T such that $T^{\wedge}ST = I$.

With reference to a fixed Hamiltonian function $H = H(x; t)$, use will be made of the differential operator ∇ which is defined by

$$(18) \quad \nabla F = F_t + H_x \cdot IF_x, \quad \text{where } F = F(x; t)$$

is a scalar function of class $C^{(1)}$; so that ∇F is a continuous scalar function in the $(2n + 1)$ -dimensional $(x; t)$ -domain.

§20. If two scalar functions F, G of the $2n$ -vector $x = (x_j)$ are of class $C^{(1)}$, one can define a continuous scalar function $(F; G)$ of x by placing

$$(19) \quad \begin{aligned} (F; G) &= F_x \cdot IG_x; \quad \text{so that} \\ (F; G) &= \sum_{i=1}^n \frac{\partial(F, G)}{\partial(x_i, x_{i+n})} = - (G; F), \end{aligned}$$

by (16). Thus, if F^1, F^2, F^3 are of class $C^{(1)}$,

$$(20_1) \quad (F^1 F^2; F^3) = (F^1; F^3) F^2 + (F^2; F^3) F^1;$$

$$(20_2) \quad (F^1 + F^2; F^3) = (F^1; F^3) + (F^2; F^3).$$

Let F^1, F^2, F^3 be of class $C^{(2)}$. Application of (19) to $F = (F^1; F^2)$, $G = F^3$ shows that $((F^1; F^2); F^3)$ is of the form

$$(21) \quad \begin{aligned} ((F^1; F^2); F^3) &= \{F^2, F^3; F^1\} - \{F^1, F^3; F^2\}, \quad \text{where} \\ \{G^1, G^2; G^3\} &= \{G^2, G^1; G^3\} \end{aligned}$$

denotes a certain trilinear expression in the partial derivatives of G^1, G^2, G^3 , and is symmetric in G^1, G^2 . Without using the explicit representation of $\{G^1, G^2; G^3\}$, one sees from (21) that

$$(22) \quad ((F^1; F^2); F^3) + ((F^2; F^3); F^1) + ((F^3; F^1); F^2) = 0.$$

Since $(F; \text{const.}) \equiv 0$ by (19), it is clear from (20₁), (20₂) that if $F = F(F^1, \dots, F^l)$ is a scalar function of a certain number, l , of independent scalar variables F^k , and if each of these F^k and a G are given as functions of class $C^{(1)}$ in $x = (x_j)$, then the relation

$$(23) \quad (F(F^1, \dots, F^m); G) = \sum_{k=1}^l (F^k; G) F_{F^k}(F^1, \dots, F^m),$$

where F_{F^k} denotes the partial derivative of $F = F(F^1, \dots, F^m)$ with respect to F^k , is an identity in x for every polynomial F ; hence, also for every F which is of class $C^{(1)}$ in its (F^1, \dots, F^m) -domain.

§21. If a fixed Hamilton function $H(x; t)$, three scalar functions

$F(x; t); F^1(x; t), F^2(x; t)$ and the partial derivatives F_t^1, F_t^2 are of class $C^{(1)}$ in a $(2n + 1)$ -dimensional $(x; t)$ -domain, (18) and (19) show that

$$(24_1) \quad \nabla F = F_t + (H; F); \quad (24_2) \quad (F^1; F^2)_t = (F_t^1; F^2) + (F^1; F_t^2)$$

are identities in this domain.

It follows that if $F(x; t), G(x; t)$ and the fixed Hamiltonian function $H(x; t)$ are of class $C^{(2)}$, then

$$(25) \quad \nabla(F; G) = (\nabla F; G) + (F; \nabla G).$$

In fact, on applying (22) to $F^1 = F, F^2 = G, F^3 = H$, and then expressing $(H; F)$ and $(G; H) = -(H; G)$ by means of (24₁), one clearly obtains

$$\begin{aligned} (H; (F; G)) &= (\nabla F - F_t; G) - (\nabla G - G_t; F) \\ &\equiv (\nabla F; G) + (F; \nabla G) - (F; G)_t, \end{aligned}$$

the last identity being implied by (20₂) and (24₂). This, when compared with (24₁), proves (25).

§22. Instead of the bilinear differential operation (19) which is applied to a pair of scalar functions F, G depending on a $2n$ -vector $x = (x_j)$, one can consider the "polar" differential operation which is applied to a $2n$ -vector $y = (y_j)$ depending on two scalar variables f, g . Assuming that $y = y(f, g)$ is of class $C^{(1)}$ in the two-dimensional (f, g) -domain under consideration, the bilinear operation in question is the one which associates with the $2n$ -vector function $y = y(f, g)$ the continuous scalar function

$$(26) \quad [f; g] = y_f \cdot I y_g; \text{ so that } [f; g] = \sum_{i=1}^n \frac{\partial(y_i, y_{n+i})}{\partial(f, g)} = -[g; f],$$

by (16). One easily verifies a relation which is dual to (21) and, corresponding to (22), implies the identity

$$(27) \quad [f^1; f^2]_{f^3} + [f^2; f^3]_{f^1} + [f^3; f^1]_{f^2} = 0$$

for any $2n$ -vector function $x = x(f^1, f^2, f^3)$ which is of class $C^{(2)}$ in three scalar variables f^1, f^2, f^3 (the subscripts f in (27) denote partial differentiations).

§23. If $F = F(x)$ is such that $(F; G) \equiv 0$ in the x -domain under consideration, F is said to be in involution with $G = G(x)$. Then G is, by (19), in involution with F . By (23), every $F = F(G)$ is in in-

volution with G . If F^1, F^2, F^3 are of class $C^{(2)}$, and if F^1 is in involution with both F^2 and F^3 , then F^1 is, by (22), in involution with $(F^2; F^3)$ also.

If l functions F^1, \dots, F^l of class $C^{(1)}$ are

(i): in involution pair by pair and

(ii): independent in the $2n$ -dimensional x -domain under consideration,

then F^1, \dots, F^l are said to form an involutory system. While (ii) and §18 (where $m = 2n$) imply only that $l \leq 2n$, conditions (i) and (ii) imply that $l \leq n$. In fact, (i) and the definition (19) require the identical vanishing of the matrix $(F_x^l \cdot IF_x^k)$ with l rows and l columns, where the $2n$ -matrix I is, by (16), skew-symmetric and non-singular. On the other hand, (ii) requires the rank l for the matrix (15 bis) with $2n$ rows and l columns. It follows, therefore, from a standard property of skew-symmetric matrices (or by a direct verification, based on the definition of I), that if $l > n$, then (ii) contradicts (i).

If (i) is replaced by the more general condition

(i bis): each of the functions $(F^i; F^k)$ of x , where $i, k = 1, \dots, l$, is a function $F = F(F^1, \dots, F^l)$ of the l given functions F^i ,

one says that F^1, \dots, F^l form a function group in the x -domain under consideration. In the case of a function group, one cannot replace $l \leq 2n$ by $l \leq n$.

If t occurs explicitly in the F , the three definitions of this article are meant to hold for every fixed t and for every x in the $(x; t)$ -domain.

§24. The definitions of §23 can be illustrated by a classical example, occurring in the problem of several bodies. To this end, choose n so as to be divisible by 3, write n for $\frac{1}{3}n$, and denote the $2 \cdot 3n = 6n$ components x_j of x by $\xi_h, \eta_h, \zeta_h; \Xi_h, H_h, Z_h$, where $h = 1, \dots, n$. Choosing n fixed scalar constants m_h and placing $\sum = \sum_{h=1}^n$, define $l = 9$ functions F^1, \dots, F^9 by

$$(29_1) \quad F_I^\xi = \sum H_h \zeta_h - \sum Z_h \eta_h, \quad F_{II}^\xi = \sum \Xi_h, \quad F_{III}^\xi = \sum m_h \xi_h - t \sum \Xi_h,$$

$$(29_2) \quad F^1 = F_I^\xi, \quad F^2 = F_I^\eta, \quad \dots, \quad F^5 = F_{II}^\eta, \quad \dots, \quad F^9 = F_{III}^\xi,$$

(29₁) defining F_ν^η, F_ν^ζ for $\nu = I, II, III$ by cyclic permutations of ξ, η, ζ and Ξ, H, Z . It will be assumed that every m_h is positive.

It is easily verified that, barring from the $6n$ -dimensional phase space a finite number of analytic hypersurfaces, not only the set

(29₂) of nine functions but also every subset of this set consists of functions which are independent in the sense of §18. Now, application of the definition (19) to pairs $F = F^r$, $G = F^s$ of functions (29₂) shows that, in view of (29₁), the matrix $((F^s; F^r))$ of 9×9 scalar functions $(F^s; F^r)$ is

$$(30) \quad ((F^s; F^r)) = \begin{pmatrix} \Phi_I & \Phi_{II} & \Phi_{III} \\ \Phi_{II} & O & M \\ \Phi_{III} & -M & O \end{pmatrix}, \quad \text{where}$$

$$\Phi_\nu = \begin{pmatrix} 0 & F_\nu^\zeta & -F_\nu^\eta \\ -F_\nu^\zeta & 0 & F_\nu^\xi \\ F_\nu^\eta & -F_\nu^\xi & 0 \end{pmatrix}$$

for $\nu = I, II, III$, while O denotes the three-rowed zero matrix, finally M the product of the positive scalar constant $\sum m_h$ and of the three-rowed unit matrix.

On comparing (30) with §23, and disregarding the hypersurfaces mentioned before, one sees that the nine functions (29₂) form a function group but not an involutory system; that the same holds for the three F_I , while the three F_{II} form an involutory system, as do the three F_{III} ; and that an F_I is in involution with an F_{II} or an F_{III} if and only if the superscripts ξ, η, ζ are identical, while an F_{II} is in involution with an F_{III} if and only if these superscripts are distinct.

§25. With reference to a reciprocal pair of phase space transformations (11₁)–(11₂), one can introduce two skew-symmetric $2n$ -matrices which are functions of the position in the $(2n+1)$ -dimensional $(y; t)$ -domain and are defined as follows: The first of these matrices, $((y_i; y_k))$, is formed by the $(2n)^2$ scalar functions $(y_i; y_k)$ which one obtains by identifying F, G in (19) with two arbitrary components $y_i = y_i(x; t)$, $y_k = y_k(x; t)$ of the $2n$ -vector (11₁), and then expressing x , as in §17, by means of (11₂) as a function of $(y; t)$; while the second matrix, $([y_i; y_k])$, is formed by the $(2n)^2$ scalar functions $[y_i; y_k]$ which one obtains by identifying the scalars f, g in (26) with two arbitrary components y_i, y_k of the $2n$ -vector y which occurs in the representation (11₂) of the $2n$ -vector x . Thus, if i and k refer to rows and columns, respectively, one sees from the definitions (19), (26) and (16), that the pair of $2n$ -matrices in question may be written as matrix products, $((y_i; y_k)) = y_x I y_x'$ and $([y_i; y_k]) = x_y' I x_y$,

where $\Gamma' = -\Gamma = \Gamma^{-1}$. It follows, therefore, from (13₂) and (15₂) that

$$(31) \quad ((y_i; y_k)) = \Gamma \Gamma' \quad \text{and} \quad ([y_i; y_k]) = (\Gamma \Gamma')^{-1};$$

so that the two matrices are transposed reciprocals of each other and are expressible in terms of the Jacobian matrix $\Gamma = y_x$.

Canonical Transformations

§26. With reference to a Hamiltonian function $H = H(x; t)$ of class $C^{(1)}$ and to a transformation (11₁)–(11₂) which satisfies the C -conditions of §17, and in accordance with the agreements (12), (13₁), (13₂), define in the $(2n + 1)$ -dimensional $(y; t)$ -domain a $2n$ -vector function $w = w^H$ by placing

$$(1) \quad w^H(y; t) \equiv w^H = I y_t + I^{-1} \Gamma \Gamma' H_y.$$

If the transformation (11₁)–(11₂) of the phase space (or, rather, the pair of vector and matrix functions $y_t(y; t)$ and $\Gamma(y; t)$ which belongs to this transformation without any reference to an H) has the property that there exists for *every* $H = H(x, t)$ a scalar function $K \equiv K^H = K^H(y; t)$ by means of which the $2n$ -vector function (1) is representable as the gradient $K_y(y; t)$ with respect to the $2n$ -vector y , then (11₁)–(11₂) is called a canonical transformation. Clearly, $K = K^H$ either does not exist or else it is uniquely determined by H up to an arbitrary additive scalar function of t alone. Correspondingly, two functions $K = K^H$ will not be considered as distinct if their difference is independent of y .

In view of the italicized word in the above definition, a K^H will exist for *certain* H also when the transformation is not canonical (e.g., $H = \text{const.}$ is such a particular H , no matter what is the transformation). The question, which are those particular H for which K exists in the case of a given non-canonical transformation, will not be discussed in what follows (the answer to this question depends on Lie's theory of function groups).

§26 bis. One is led to the $2n$ -vector function $w^H(y; t)$, and then to the notion of a canonical transformation, if one subjects the operation (17) to an arbitrary transformation (11₁)–(11₂), where it is understood that the $2n$ -vector $x' + I H_x(x; t)$ is not a function of the position in the $(2n + 1)$ -dimensional $(x; t)$ -domain, since it is defined only with reference to an arbitrarily given path $(x(t); t)$ of class $C^{(1)}$ in this domain.

First, it is clear from (15₃) and (12) that (11₁) transforms $x' + IH_x$ into the sum of $\Gamma^{-1}y' - \Gamma^{-1}y_t$ and $I\Gamma'H_v$; so that, since $I = -I^{-1}$ by (16),

$$(2) \quad x' + IH_x = \Gamma^{-1}\{y' + Ily_t + II^{-1}\Gamma I\Gamma'H_v\} \equiv \Gamma^{-1}\{y' + Iw^H\},$$

by the above definition (1). This means that, whether the transformation (11₁) is canonical or not, its non-singular Jacobian matrix (13₂) transforms the vector function $x' + IH_x$ of t into the vector function $y' + Iw^H$ of t along any path of class $C^{(1)}$. It follows that the transformation (11₁) is canonical if and only if the vector $x' + IH_x$ is transformed in case of an *arbitrary* Hamiltonian function $H(x; t)$ and along an *arbitrary* path into a vector of the same form; that is, into $y' + IK_v$, where the new Hamiltonian function, $K \equiv K(y; t) = K^H$, depends on H but not on the choice of the path.

§27. It will be proved that a transformation $y = y(x; t)$, $x = x(y; t)$ of the type considered in §17 is a canonical transformation if and only if there exists a scalar μ which is a constant in the $(2n + 1)$ -dimensional domain under consideration and is such that the matrix relation

$$(3) \quad \Gamma I \Gamma' = \mu I \quad \text{where} \quad \Gamma \equiv \Gamma(y; t) = y_x, \\ (I^{-1} = I' = -I, \det I = +1),$$

is an identity in this domain. According to (31), §25, one can express this condition in terms of either of the $2n$ -matrices $((y_i; y_k))$, $([y_i; y_k])$.

Application of the multiplication theorem of determinants to (3) shows that the absolute value of the constant μ is uniquely determined by the Jacobian $\det \Gamma$ ($\neq 0$), since

$$(4) \quad (\det \Gamma)^2 = \mu^{2n}; \quad \text{so that} \quad 0 \neq |\det \Gamma(y; t)| = |\mu|^n = \text{const.}$$

The Hamiltonian function $K = K(y; t)$ into which a canonical transformation sends an arbitrary Hamiltonian function $H(x; t)$ will turn out to be

$$(5) \quad K = \mu H + R,$$

where $H(x; t)$ is thought of as expressed by means of $x = x(y; t)$ as a function of $(y; t)$, and $R = R(y; t)$ denotes a scalar function for which $R_t(y; t)$ is of class $C^{(1)}$.

Finally, it will be shown that R and the $2n$ -vector (13₁), §17, are connected by the identity

$$(6) \quad Iy_t = R_y, \quad \text{where} \quad y_t = y_t(y; t), \quad R = R(y; t).$$

This implies that, in (5), not only μ but also R depends merely on the canonical transformation $y = y(x; t)$ and not on the choice of H . Actually, $R = R(y; t)$ follows from (6) by a quadrature in the y -domain for a fixed t ; so that an additive function of t alone remains undetermined. This agrees, in view of (5), with §26. Correspondingly, two R are to be considered as identical if their difference is independent of y .

In view of (3) and (5), the scalars μ and $R(y; t)$ which belong to a canonical transformation $y = y(x; t)$ will be called its multiplier and its remainder function, respectively.

The proof of the statements of the present article will be supplied in §28–§30.

§28. First, the lemma formulated at the end of §4 can easily be applied to (1), by identifying a, A, f_x, m with $y_t, I^{-1}\Gamma I\Gamma^*, H_y, 2n$, respectively, and keeping t fixed. It then follows from that lemma that a given transformation of the type considered in §17 will make (1), §26 the gradient, $K_y = K_y(y; t)$, of a suitable $K = K^H$ for every H if and only if Iy_t is the gradient, R_y , of a suitable $R = R(y; t)$, and $I^{-1}\Gamma I\Gamma^*$ is the product of the unit $2n$ -matrix and of a scalar μ which is independent of y , i.e., which depends on the parameter t alone. In other words, the transformation is canonical if and only if there exist suitable $R = R(y; t)$ and $\mu = \mu(t)$ satisfying (6) and (3).

Finally, substitution of (6), (3) into (1) gives $w^H = R_y + I^{-1}\mu IH_y$, where $w^H = K_y$ and $\mu_y \equiv 0$. Hence, $K_y = (R + \mu H)_y$; and so (5) follows by neglecting an arbitrary additive function of t alone.

§29. The criterion proved in §28 seems to be at variance with the criterion announced in §27, since, while either of these criteria is both necessary and sufficient for a transformation which is canonical, §28 does, and §27 does not, allow μ to depend on t .

The answer is that one cannot find for an arbitrarily given pair R, μ a transformation $x = x(y; t)$ satisfying (6), (3). In fact, $y_t = y_t(y; t)$ and $\Gamma = \Gamma(y; t) \equiv y_x$; so that (6), (3) represent quite a complicated system of partial differential equations for the $2n$ -vector function $x = x(y; t)$.

Now, §30 will imply that $\mu = \text{const.}$ is an integrability condition of these partial differential equations (so that §27 follows from §28). Since (3) holds, by §28, for a suitable $\mu = \mu(t)$, and since (3) implies (4), it will be sufficient to prove that $(\det \Gamma)^2$ cannot depend on t .

Since $\det I = +1$ implies that also $\det (\Gamma' I \Gamma) = (\det \Gamma)^2$, it will be sufficient to prove that the matrix $\Gamma' I \Gamma$ cannot depend on t .

§30. To this end, it will be shown that for an arbitrary transformation $x = x(y; t)$, which need not satisfy (3) with a $\mu = \mu(t)$ and not even with a $\mu = \mu(y; t)$, there does or does not exist a scalar $R = R(y; t)$ satisfying (6) according as the matrix $\Gamma' I \Gamma$, defined by $y_x = \Gamma \equiv \Gamma(y; t)$ as a function of y and t , is or is not independent of t .

First, it is easily verified from §17 that, I being the matrix (16), §19,

$$(7) \quad (I y_t)_y = I(y_t)_y; \quad \text{so that} \quad (I y_t)_y = I \Gamma_t \Gamma^{-1},$$

by (15₁), §17. Hence, the matrix $(I y_t)_y$ is symmetric if and only if

$$(8) \quad I \Gamma_t \Gamma^{-1} = (I \Gamma_t \Gamma^{-1})', \quad \text{i.e.,} \quad \Gamma' I \Gamma_t + \Gamma'_t I \Gamma = 0; \quad (I' = -I = I^{-1}).$$

Since $I = \text{const.}$, this can be written as $(\Gamma' I \Gamma)_t = 0$. It follows that $(I y_t)_y$ is a symmetric matrix if and only if the $2n$ -matrix function $\Gamma' I \Gamma$ of y and t is independent of t . But $(I y_t)_y$ is, by the beginning of §3, a symmetric matrix if and only if the vector $I y_t$ is a gradient, i.e., if and only if there exists an $R = R(y; t)$ such that (6) is an identity in y for every fixed t .

This completes the proof of fact stated at the beginning of this article. Hence, §29 shows that the proof of the statements of §27 is now complete.

§31. It is clear from the definition (§26) of a canonical transformation that the set of all canonical transformations defined on a common $(2n + 1)$ -dimensional domain is a group. The composition rule of the Jacobian matrices Γ , remainder functions R and multipliers μ is that if $\Gamma_1, R_1, \mu_1; \Gamma_2, R_2, \mu_2$ belong to two canonical transformations and Γ, R, μ to the canonical transformation which is obtained by applying the second of these transformations after the first, then

$$(9_1) \quad \Gamma = \Gamma_2 \Gamma_1; \quad (9_2) \quad \mu = \mu_1 \mu_2; \quad (9_3) \quad R = \mu_2 R_1 + R_2.$$

This is easily verified from (3) and (6). It is also seen from (3) and (6) that if E denotes the unit $2n$ -matrix, $\Gamma \equiv E, \mu = 1, R \equiv 0$ belong to the identical transformation $y = x$. It follows, therefore, from (9₁), (9₃), (9₂) that if Γ, R, μ belong to a canonical transformation, then

(10) Γ^{-1} , $-\mu^{-1}R$, μ^{-1} belong to the inverse transformation.

§31 bis. It may be mentioned that a transformation is canonical if and only if

(11) $\Gamma' I \Gamma = \mu I$, where $\Gamma \equiv \Gamma(y; t) = y_x$, $\mu = \text{const.} \neq 0$, i.e., that (3) is equivalent to (11). In fact, if $\Gamma I \Gamma' = \mu I \neq 0$, then $\Gamma' = \mu I^{-1} \Gamma^{-1} I$, since $I^{-1} = -I$; so that $\Gamma' = \mu I \Gamma^{-1} I^{-1}$, and so $\Gamma' I \Gamma = \mu I$.

§32. It will be proved in §62–§62 bis that (3) implies

(12) $\det \Gamma(y; t) = \mu^n$, ($\mu = \text{const.} \neq 0$),

a relation which is, in case of an odd degree of freedom n , sharper than (4).

§33. If x^ν and y^ν , where $\nu = \text{I, II}$, are four $2n_\nu$ -vectors, let $x^{\text{I II}}$ and $y^{\text{I II}}$ denote the $2(n_{\text{I}} + n_{\text{II}})$ -vectors obtained by uniting the components of x^{I} , x^{II} and y^{I} , y^{II} . If both component transformations $x^\nu = x^\nu(y^\nu; t)$ are canonical, μ^ν and $R^\nu = R^\nu(y^\nu; t)$ denote the respective multipliers and characteristic functions, and if $\mu^{\text{I}} = \mu^{\text{II}}$, then it is clear from §27 that the resulting transformation $x^{\text{I II}} = x^{\text{I II}}(y^{\text{I II}}; t)$ is again canonical and has the multiplier $\mu^{\text{I}} = \mu^{\text{II}}$ and the remainder function $R^{\text{I}} + R^{\text{II}}$.

§34. A canonical transformation $y = y(x; t)$ or $x = x(y; t)$ is said to be completely canonical if it transforms every Hamiltonian function $H(x; t)$ into a Hamiltonian function $K(y; t)$ which is identical with $H(x; t)$ in virtue of $y = y(x; t)$; so that, for every H ,

(13) $K(y; t) = H(x(y; t); t)$, i.e., $\mu = 1$, $R(y; t) = 0$, ($\det \Gamma = 1$).

Cf. (5), (12). Clearly, these transformations form a subgroup of the group mentioned in §31.

§35. Another subgroup is obtained by considering those canonical transformations $x = x(y; t)$ which are conservative in the sense defined at the end of §18; so that $x = x(y)$. It is clear from (6) that $R(y; t) = 0$ holds for the transformations of this subgroup also; so that (5) reduces to $K = \mu H$. It follows, therefore, from (13) that a conservative canonical transformation is completely canonical if and only if its multiplier is $+1$.

§35 bis. If $F(x)$, $G(x)$ are two scalar functions of class $C^{(1)}$, let the definition (19), §20 be written as $(F; G)^x = F_x \cdot I G_x$, the super-

script emphasizing the dependence of $(F; G)$ on the coordinate system x . If $y = y(x)$ is another coordinate system, then $(F; G)^x = (\Gamma F_y) \cdot (\Gamma G_y)$, by (12), §17; so that $(F; G)^x = F_y \cdot \Gamma' \Gamma G_y$, by §1. Hence, if (11) is satisfied (and, when F and G are unspecified, *only* if (11) is satisfied), one has $(F; G)^x = \mu(F; G)^y$. Accordingly, those conservative transformations $y = y(x)$ which are canonical are characterized by the property that they leave $(F; G)$ relative-invariant* for arbitrary F and G , where the adjective "relative" refers to the appearance of an arbitrary constant factor μ (so that $\mu = 1$ in case of absolute invariance).

§36. If $x = x(y; t)$ is a canonical transformation and t_0 denotes some fixed value of t , the conservative transformation $x = x(y; t_0)$ is canonical. This is clear from the criterion (3), where $\mu = \text{const.}$ It is also seen that if it is known only that $x = x(y; t)$ is such as to make the conservative transformation $x = x(y; t_0)$ canonical for every fixed t_0 , then $x = x(y; t)$ need not be a canonical transformation, since then nothing guarantees that μ is independent of t , i.e., that the integrability condition (8) of (6) is satisfied. All that follows from §30 and §28 is that if a transformation $x = x(y; t)$ satisfies

(i): the condition (3) for every y at a fixed $t = t_0$, and

(ii): the gradient condition (6) for every y at every t ,

then it satisfies (3) for every y at every t and is a canonical transformation.

§37. Consider, finally, the subgroup of those canonical transformations which are (homogeneous and) linear in the $2n$ coordinates of the phase space, i.e., for which $y = \Gamma x$, where $\Gamma = \Gamma(t)$ is a given non-singular $2n$ -matrix on some t -interval. For this subgroup, (3) and (6) respectively reduce to

$$(14_1) \quad \Gamma \Gamma' = \mu I, \text{ where } \Gamma = \Gamma(t), \mu = \text{const.} \neq 0;$$

$$(14_2) \quad R = \frac{1}{2} y \cdot \Gamma' \Gamma^{-1} y$$

In fact, the Jacobian matrix, $y_x = \Gamma \equiv \Gamma(y; t)$, of $y = \Gamma(t)x$ is $\Gamma(t)$. Hence, $\Gamma_t = d\Gamma/dt \equiv \Gamma'$, and so $y_t = \Gamma' x$, where $x = \Gamma^{-1} y$; so that the remainder function $R = R(y; t)$ is, in view of (6), the quadratic form (14₂) in the $2n$ components y_j of y (the matrix of the form (14₂) is a function of t alone and is, by (8), necessarily sym-

* As a consequence, the notion of involutory function pairs (§23) is canonically invariant.

metric, if the condition (14₁) for a canonical transformation $y = \Gamma(t)x$ is satisfied).

§38. If, for instance, the $2n$ -matrix $\Gamma(t)$ is obtained by repeating an orthogonal n -matrix $P(t)$ along the principal diagonal, i.e., if

$$(15_1) \quad \Gamma(t) = \begin{pmatrix} P & (0) \\ (0) & P \end{pmatrix}, \text{ where } P = P(t), P' = P^{-1}, \text{ then } \Gamma I \Gamma' = I,$$

by the definition of I ; so that (14₁) is satisfied by $\mu = 1$, while (14₂) becomes

$$(15_2) \quad 2R(y; t) = u \cdot P' P' v - v \cdot P' P' u, \quad (P = P(t) = P'^{-1}),$$

if $u = (u_i)$, $v = (v_i)$ denote the n -vectors defined by (10), §17.

For instance, if n is even and $P(t)$ is the particular orthogonal n -matrix obtained by repeating, $\frac{1}{2}n$ times along the principal diagonal, an orthogonal 2-matrix

$$(16) \quad \Phi(t) = \begin{pmatrix} \cos \phi(t) & -\sin \phi(t) \\ \sin \phi(t) & \cos \phi(t) \end{pmatrix}, \quad \text{then}$$

$$R(y; t) = \phi'(t) \sum_{k=1}^{\frac{1}{2}n} (\xi_k H_k - \eta_k \Xi_k),$$

if in (15₂) one puts $u_{2k-1} = \Xi_k$, $u_{2k} = H_k$; $v_{2k-1} = \xi_k$, $v_{2k} = \eta_k$.

If $\Gamma(t)$ is a 2-matrix, so that $n = 1$, i.e.,

$$(17_1) \quad \Gamma(t) = \begin{pmatrix} a(t) & b(t) \\ c(t) & d(t) \end{pmatrix}, \quad \text{then} \quad \det \Gamma(t) = \mu = \text{const.} \neq 0$$

is equivalent to (14₁), while (14₂) reduces, if $y \equiv \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} u \\ v \end{pmatrix}$, to

$$(17_2) \quad 2\mu R = D_{cd}u^2 + (D_{bc} - D_{ad})uv + D_{ab}v^2, \text{ where } D_{fg} = f'g - g'f.$$

Canonical Transformations and Pfaffians

§39. If an arbitrary phase space transformation $y = y(x; t)$; $x = x(y; t)$ of the pair of $2n$ -vectors $x = (x_i)$; $y = (y_i)$ is expressed in terms of the four n -vectors $p = (p_i)$, $q = (q_i)$; $u = (u_i)$, $v = (v_i)$ formed by the momenta p_i , u_i and the coordinates q_i , v_i , then one has to write

$$(1_1) \quad \begin{aligned} u &= u(p, q; t) \\ v &= v(p, q; t); \end{aligned}$$

$$(1_2) \quad \begin{aligned} p &= p(u, v; t) \\ q &= q(u, v; t); \end{aligned}$$

$$(2) \quad x = \begin{pmatrix} p \\ q \end{pmatrix}, \quad y = \begin{pmatrix} u \\ v \end{pmatrix}; \quad (3) \quad \Gamma(u, v; t) = y_x = \begin{pmatrix} u_p & u_q \\ v_p & v_q \end{pmatrix};$$

$$(4) \quad \tau I = \begin{pmatrix} (0) & (e_k^i) \\ - (e_k^i) & (0) \end{pmatrix} = -I = -I^{-1}.$$

If the phase space transformation (1₁) is of class $C^{(1)}$ and such that both n -vector functions u_t, v_t are of class $C^{(1)}$ in the $(2n + 1)$ -dimensional domain under consideration, the criterion (3), §27 states that the transformation (1₁)–(1₂) is canonical if and only if the three n -vector relations

$$(5) \quad u_p u_q^{\backslash} = u_q u_p^{\backslash}; \quad v_q v_p^{\backslash} = v_p v_q^{\backslash}; \quad u_p v_q^{\backslash} - u_q v_p^{\backslash} = \mu(e_k^i),$$

where $\mu = \text{const.} \neq 0$, are identities in $(u, v; t)$ in virtue of (1₂). The first two of these conditions can be expressed by saying that the products $u_p u_q^{\backslash}$ and $v_q v_p^{\backslash}$ are symmetric matrices. Notice that the Jacobian n -matrices which constitute the Jacobian $2n$ -matrix (3) can have vanishing determinants, although $\det \Gamma \neq 0$.

If (5) is satisfied, then, by (5), §27 and (6), §27,

$$(6) \quad K = \mu H + R; \quad (7) \quad v_t = R_u(u, v; t), \quad -u_t = R_v(u, v; t),$$

where v_t, u_t are obtained by differentiating (1₁) with respect to t at fixed p, q , and then expressing p, q by means of (1₂) in terms of $(u, v; t)$; cf. (13₁), §17.

§40. A transformation (1₁)–(1₂) will be called binary if it belongs to the degree of freedom $n = 1$; so that p, q, u, v are scalars. Thus, the matrices u_p, u_q, \dots are scalars, hence commutable and such that the sign of transposition can be omitted. Consequently, the first two of the three conditions (5) reduce to $0 = 0$, while the third is easily seen to be equivalent to

$$(8) \quad \partial(u, v)/\partial(p, q) \equiv \mu = \text{const.} \neq 0, \quad \text{where} \\ u = u(p, q; t), \quad v = v(p, q; t).$$

Thus, a binary transformation (1₂) is canonical if and only if its Jacobian matrix is of constant determinant ($\neq 0$).

§41. It is clear from §40 and §35 that a conservative binary transformation is completely canonical if and only if

$$(9) \quad \partial(u, v)/\partial(p, q) \equiv +1, \quad \text{where} \quad u = u(p, q), \quad v = v(p, q),$$

i.e., if and only if the mapping (of class $C^{(1)}$) which sends a domain of

the (p, q) -plane into a domain of the (u, v) -plane is area and* orientation preserving.

For instance, condition (9) is satisfied, if $p > 0$, by

$$(10) \quad u = \sqrt{2p} \cos q, \quad v = \sqrt{2p} \sin q, \quad \text{where} \quad \sqrt{2p} \geq 0.$$

On the other hand, the introduction $u = p \cos q, v = p \sin q$ of polar coordinates into a Cartesian phase plane (u, v) is not a canonical transformation, since the Jacobian (8) becomes $p \neq \text{const.}$

§42. From now on, the number n of the components of each of the vectors $p = (p_i), q = (q_i)$ occurring in (2) will again be arbitrary.

If $n = 1$, then either of the transformations $u = \pm q, v = \mp p$ is, by §41, completely canonical. Hence, the same holds, by §33, for any n .

If u_1, \dots, u_n is any permutation of p_1, \dots, p_n , and if v_1, \dots, v_n is the same permutation of q_1, \dots, q_n , then $u = p, v = q$ is a completely canonical transformation. This is clear from §33 (and also follows by choosing the n -matrix P in (15₁), §38 so as to contain 1 as an element of each of its rows).

§42 bis. Also the addition of arbitrary constants to the p_i, q_i is a canonical transformation of multiplier $\mu = 1$, since $\Gamma = y_x$ then is the unit matrix.

§43. If there is given no relation between t and the pair $x = (x_j), y = (y_j)$ of $2n$ -vectors, then

$$(11) \quad \omega = 2R dt + \mu x \cdot Idx - y \cdot Idy \quad [\text{cf. (4)}]$$

is, for arbitrarily given scalar functions R, μ of $(t; x, y)$, a scalar Pfaffian in $4n + 1$ independent variables. Suppose that there is given, for every fixed t , a relation between x and y in the form

$$(12) \quad F(t; x, y) = 0, \quad \text{where} \quad F = (F_j)$$

is a $2n$ -vector and the F_j are, for arbitrary fixed t , independent in the sense of §18. Then the Pfaffian (11) in $4n + 1$ variables becomes in virtue of (12) a Pfaffian in $2n + 1$ variables. It will be assumed that $F_i(t; x, y)$ is of class $C^{(1)}$ in the $(4n + 1)$ -dimensional domain. Then (12) is, for every fixed t , an implicit definition of a locally topological mapping of the $2n$ -dimensional phase spaces x, y on each other; and the mapping functions and their partial deriva-

* If it is only area preserving, the Jacobian (9), i.e. μ , is -1 .

tives with respect to t are of class $C^{(1)}$ in the respective $(2n + 1)$ -dimensional domains.

It will be shown that the mapping which is implicitly defined by (12) is a canonical transformation if and only if there exist a constant $\mu \neq 0$ and a scalar function R such that the Pfaffian in $2n + 1$ variables to which the Pfaffian (11) in $4n + 1$ variables reduces in virtue of (12) is a complete differential.*

Notice that the property of being a complete differential is an invariant property. In fact, this property of a Pfaffian is characterized by the symmetry of the Jacobian matrix of the (covariant) coefficient vector function of the Pfaffian, i.e., by the identical vanishing of the curl (cf. the beginning of §3). Hence, the statement follows from the fact that the curl is a tensor.†

Due to the invariance just mentioned, it will be sufficient to consider (11) on the assumption that (12) is given in the explicit form $y = y(x; t)$.

The calculations will always use the fact that $a \cdot Cb = b \cdot C'a$, by §1; and that $I' = -I = I^{-1}$.

§44. First, it is clear from (15₃), §17, that, whether the transformation $y = y(x; t)$, implicitly defined by (12), is canonical or not, the Pfaffian (11) in $4n + 1$ variables reduces in virtue of (12) to

$$(13) \quad \omega = Tdt + X \cdot dx;$$

$$(14_1) \quad T = 2R - y \cdot Iy_t; \quad (14_2) \quad X = -\mu Ix + \Gamma' Iy,$$

where $R = R(t; x, y)$, $\mu = \mu(t; x, y)$ are scalar functions given with (11), while the scalar T , defined by (14₁), and the (covariant) $2n$ -vector X , defined by (14₂), are thought of as expressed by means of $y = y(x; t)$ as functions of $(x; t)$; so that (13) is a Pfaffian in $2n + 1$ independent variables $x_1, \dots, x_{2n}; t$, the dot denoting scalar multiplication of X and dx .

By the beginning of §3, the Pfaffian (13) is a complete differential if and only if the $(2n + 1)$ -vector formed by the scalar T and the $2n$ components of X has, with respect to the $(2n + 1)$ -vector formed by

* This characteristic property of the canonical transformations is equivalent to Lie's definition of a contact transformation, provided that t is considered as an additional coordinate; a coordinate which can be transformed in the same way as the $2n$ coordinates of the phase space (cf., e.g., §9 bis).

† This elementary fact cannot be expressed by saying that the curl is the difference of two covariant derivatives, since this manner of speaking presupposes a differential geometry. The verification is, however, straightforward in every case.

the scalar t and the $2n$ components of x , a Jacobian matrix which is symmetric for every $(x; t)$. Since this symmetry condition is expressed by the pair of conditions

$$(15_1) \quad X_t = T_x;$$

$$(15_2) \quad X_x = X'_x,$$

the criterion announced in §43 will be proved if one shows that (15₁) and (15₂) together are equivalent to the criterion (3) of §27, where $\mu = \text{const.}$ Hence, it is clear from §28 and §30 that it is sufficient to prove that

(i): if μ in (11) is independent of t , then the vector condition (15₁) is equivalent to the existence of an R which satisfies (6), §27;

(ii): if $\mu = \mu(t)$ in (11), then the matrix condition (15₂) is equivalent to (3), §27, i.e., to $\Gamma'I\Gamma = \mu I$ (cf. §31 bis).

§44 bis. First, the gradient, $(y \cdot Iy_t)_x$, of $y \cdot Iy_t \equiv -y_t \cdot Iy$ is obviously $(y_x)'Iy_t - (y_t)_x Iy$. But $(y_t)_x \equiv (y_x)_t$; so that, since $y_x = \Gamma$ by (3), it follows from (14₁) that

$$T_x = 2R_x - \Gamma'Iy_t + \Gamma_t'Iy, \quad \text{where} \quad R_x = \Gamma'R_y, \quad \text{by (12), §17.}$$

But $x_t (\neq x')$ is identically 0, since x and t form the $(2n + 1)$ -dimensional domain of the independent variables. Hence, it is seen from (14₂) that if μ_t is identically 0, i.e., if $\mu = \mu(x)$, then (15₁) is equivalent to $(\Gamma'Iy)_t = T_x$, and so, by the above representation of T_x , to

$$(\Gamma'Iy)_t = 2\Gamma'R_y - \Gamma'Iy_t + \Gamma_t'Iy.$$

Since $(\Gamma'Iy)_t \equiv \Gamma_t'Iy + \Gamma'Iy_t$, the last relation is equivalent to $2\Gamma'Iy_t = 2\Gamma'R_y$, i.e., to $Iy_t = R_y$. This proves (i), §44.

Next, if μ_x is identically 0, i.e., if $\mu = \mu(t)$, then $X_x = -\mu I + (\Gamma'Iy)_x$, by (14₂); so that (15₂) then is equivalent to

$$\frac{1}{2} \{ (\Gamma'Iy)_x - (\Gamma'Iy)_x' \} = \mu I, \quad \text{since} \quad I' = -I.$$

But Γ is defined as the Jacobian matrix y_x of the point transformation $y = y(x; t)$ at a fixed t , while the $2n$ -matrix $\{ \}$ occurring in the equivalent formulation, $\frac{1}{2} \{ \} = \mu I$, of the assumptions (15₂) represents the curl of the $2n$ -vector function $\Gamma'Iy$ of the $2n$ -vector x at a fixed t . Since a curl is transformed by a point transformation as a tensor (§43), the proof of (ii), §44, is complete.

This proves the Pfaffian criterion announced in §43.

§45. Using the notations (2) of §39, one can write (12) as

$$(16) \quad F_j(t; p, q, u, v) = 0, \quad \text{where} \quad j = 1, \dots, 2n,$$

while (11) becomes, if $a \cdot db$ denotes $\sum_{i=1}^n a_i db_i$,

$$(17) \quad \frac{1}{2}\omega = Rdt + \mu \frac{1}{2}(p \cdot dq - q \cdot dp) - \frac{1}{2}(u \cdot dv - v \cdot du); \quad \text{cf. (4).}$$

Hence, by the criterion of §43, the transformation (1₁)–(1₂) which is implicitly defined by (16) is a canonical transformation if and only if there exist an $R = R(t; p, q, u, v)$ and a $\mu = \text{const.} \neq 0$ for which the Pfaffian (17) becomes a complete differential in virtue of (16).

This criterion remains unchanged if one adds to the Pfaffian (17) the complete differential

$$df = f_t dt + f_p \cdot dp + f_q \cdot dq + f_u \cdot du + f_v \cdot dv$$

of any scalar $f = f(t; p, q, u, v)$. Choosing, in particular,

$$f = \frac{1}{2}\mu p \cdot q \pm \frac{1}{2}u \cdot v,$$

where $\mu = \text{const.}$, one sees that the criterion remains valid if (17) is replaced by either of the Pfaffians ω_+ , ω_- , where

$$(18_1) \quad \omega_+ = Rdt + \mu p \cdot dq + v \cdot du; \quad (18_2) \quad \omega_- = Rdt + \mu p \cdot dq - u \cdot dv.$$

§45 bis. Since the criteria of §27, §28, §36, §43, §45 for a canonical transformation are all equivalent, one can tell only with a given application in view, which of these criteria is the most convenient. The Pfaffian criteria are prepared, of course, for cases where the transformation is implicitly defined by means of $2n$ independent relations (16) between the $4n + 1$ variables $t; p_i, q_i, u_i, v_i$.

§46. Let $S = S(t; q; u)$ be any scalar function of class $C^{(2)}$ in a $(2n + 1)$ -dimensional $(t; q; u)$ -domain, and suppose that, in this domain, the n -rowed “polar Hessian” matrix $(S_q)_u$ is non-singular, i.e., that

$$(19) \quad \det (S_{q_i u_k}) \neq 0, \quad \text{where} \quad S_{q_i u_k} = S_{u_k q_i}(t; q; u);$$

$$i, k = 1, \dots, n.$$

Then the pair of n -vector equations

$$(20) \quad p - S_q(t; q; u) = 0, \quad v - S_u(t; q; u) = 0$$

defines a canonical transformation (1₁)–(1₂); furthermore,

$$(21) \quad \mu = 1, \quad R = S_t; \quad \text{so that} \quad K = H + S_t, \quad \text{by (6).}$$

In order to prove this, one has to identify (16) with (20); so that

$$F_i \equiv p_i - S_{q_i}, \quad F_{i+n} \equiv v_i - S_{u_i}, \quad \text{where} \quad i = 1, \dots, n \quad \text{and} \quad S = S(t; q; u).$$

Thus, the Jacobian of $F_1, \dots, F_n, F_{n+1}, \dots, F_{2n}$ with respect to $p_1, \dots, p_n, q_1, \dots, q_n$ reduces to the n -rowed Jacobian $(-1)^n \det (S_{q_i u_k})$, and so it is, by (19), distinct from 0. Hence, it is clear from the corresponding remarks of §43, that (20) implicitly defines a transformation (1₁)–(1₂). In order to see that this transformation is canonical and has S_t and 1 as remainder function and multiplier, respectively, it is sufficient to observe that the Pfaffian (18₁) becomes in virtue of (20) a complete differential, $dS(t; q; u)$, if one chooses $R = S_t, \mu = 1$.

One must not make, however, the mistake of believing* that there exists for every canonical transformation of multiplier $\mu = 1$ an $S = S(t; q; u)$ by means of which the transformation is representable in the form (20). It is true, by §45, that a transformation which is defined implicitly is canonical with multiplier $\mu = 1$ if and only if the Pfaffian $Rdt + p \cdot dq + v \cdot du$ becomes a complete differential. But this does not imply the explicit existence of a function $S = S(t; q; u)$ for which (19) is satisfied and $S_t = R, S_q = p, S_u = v$. For instance, $p = v, q = -u$ is a canonical transformation of multiplier $\mu = 1$, although there does not exist an $S(t; q; u)$ satisfying (20).

On the other hand, it is clear from §42 that one can start with an S which contains, instead of the u_i and the q_i , any $2n$ of the $4n$ variables p_i, q_i, u_i, v_i ; e.g., an arbitrary pair selected from p, q, u, v .

For instance, if $S = S(t; q; v)$, one has to replace (19) by

$$(22) \quad \det (S_{q_i v_k}) \neq 0, \quad \text{where} \quad S_{q_i v_k} = S_{v_k q_i}(t; q; v);$$

$$i, k = 1, \dots, n.$$

Then, if (18₂) is used instead of (18₁), it follows that

$$(23) \quad p - S_q(t; q; v) = 0, \quad u + S_v(t; q; v) = 0$$

defines a canonical transformation for which (21) is again valid.

According to (21) and §34, these transformations are completely canonical if and only if t does not occur in S .

Extended Coordinate Transformations

§47. Consider, as in §10, a mapping of two n -dimensional configuration spaces $q, \bar{q} \equiv v$ on each other; so that

$$(1) \quad v = v(q; t); \quad (2) \quad \det J \neq 0, \quad \text{where} \quad J = v_q = J(q; t).$$

* This mistake is made, in particular, by those text-books of quantum theory which claim a simplification of the theory of canonical transformations.

The n -vector function $v(q; t)$ will be supposed to be of class $C^{(2)}$ in the $(n + 1)$ -dimensional $(q; t)$ -domain.

One can extend the coordinate transformation (1) in various ways to transformations (1₁)–(1₂), §39, of the $2n$ -dimensional phase spaces (2), §39, the choice of $u = u(p, q; t)$ being practically unrestricted. It turns out that, among these extensions of a given coordinate transformation (1) to phase space transformations (1₁)–(1₂), §39, there always exist canonical transformations. This may be inferred, for instance, from the criterion (17), §45, which also shows that the canonical mate $u = u(p, q; t)$ of the given coordinate transformation $v = v(q; t)$ is not uniquely determined by the latter; $\mu = \text{const.} \neq 0$ and R being unrestricted.

§48. Actually, one can choose a canonical extension

$$(3) \quad u = u(p, q; t), \quad v = v(q; t)$$

of (1) in such a way that μ becomes $+1$ and $u = u(p, q; t)$ homogeneous and linear in the components of p , namely, $u = J^{-1}p$; cf. (2).

To this end, one can choose

$$(4) \quad \mu = 1, R = v_t \cdot J^{-1}p; \text{ so that } R = R(p, q; t), \text{ by (1)–(2).}$$

Then the resulting canonical transformation (3), which will be called *the canonical extension of the given coordinate transformation (1)*, is given by

$$(5_1) \quad \begin{aligned} u &= J^{-1}p \\ v &= v(q; t); \end{aligned} \quad (5_2) \quad \begin{aligned} J &= J(q; t) \\ J &= v_{q_i}, \det J \neq 0; \end{aligned} \quad (5_3) \quad \Gamma = \begin{pmatrix} J^{-1} (J^{-1}p)_a \\ (0) & J \end{pmatrix}.$$

In fact, $dv = Jdq + v_t dt$, by (1)–(2). Since $(Aa) \cdot (Bb) = a \cdot A^* Bb$ (cf. §1), it follows that, if $u = J^{-1}p$, then $u \cdot dv = p \cdot dq + v_t \cdot J^{-1}p dt$. On substituting this and (4) into (18₂), §45, one sees that ω_- becomes a complete differential, namely $\equiv 0$. Hence, (5₁) is a canonical transformation belonging to (4). Finally, (5₃) is clear from (5₁)–(5₂) in view of the notations (1₁)–(3), §39.

According to (5₂), the extension (5₁) of $v = v(q; t)$ can be obtained by considering the momenta p_1, \dots, p_n as the components of a covariant vector in the space of the coordinates q_1, \dots, q_n at every fixed t .

§49. Suppose, in particular, that the given coordinate transformation (1) is conservative, $v = v(q)$. Then (5₁)–(5₂) reduce to

$$(6) \quad u = J^{-1}p, \quad v = v(q), \quad \text{where} \quad J = v_q = J(q), \quad \det J \neq 0;$$

while (4) becomes $\mu = 1$, $R \equiv 0$, since $v_t \equiv 0$. Thus, the canonical extension (6) of every conservative coordinate transformation $v = v(q)$ is conservative and completely canonical (cf. §34).*

§49 bis. It is obvious from the definition of a tensor, that if a transformation of a space is involutory,† then so is the transformation of the tensors of the space. Since (6) defines, for every coordinate transformation $v = v(q)$, the momenta as the components of a covariant vector in the configuration space, it follows that the canonical extension of every involutory coordinate transformation $v = v(q)$ is involutory.

§50. Suppose, for instance, that $v = v(q)$ is given as the involutory operation of a transformation by reciprocal radii; so that $v = q/|q|^2$, where $|q| = \sqrt{q \cdot q} > 0$. Then, $r_i r_k$ denoting the product of two components of an n -vector $r = (r_i)$, one has

$$(7_1) \quad J = (|v|^2 e_{ik} - 2v_i v_k); \quad (7_2) \quad J^{-1} = (|q|^2 e_{ik} - 2q_i q_k); \quad (7_3) \quad J = J^{\wedge},$$

where (e_{ik}) is the unit n -matrix. In fact, partial differentiations of $v = q/|q|^2$ show that the Jacobian matrix $J = v_q$ is the sum of the matrices $-2|q|^{-4}(q_i q_k)$ and $|q|^{-2}(e_{ik})$. Hence, (7₁) follows by using $v_l = q_l/|q|^2$ for $l = i, k$ and noting that $|v|^2 = |q|^{-2}$. And (7₂) follows from (7₁) without any calculation, by observing that $v = q/|q|^2$ is an involutory transformation.

According to (6) and (7₂), the canonical extension of $v = q/|q|^2$ is

$$(8_1) \quad v = q/|q|^2, \quad u = |q|^2 p - 2\sigma q, \quad \text{where} \quad \sigma = p \cdot q \quad (q \neq 0).$$

Since $v = q/|q|^2$ is involutory, so is (8₁), by §49 bis. Hence, the inverse of (8₁) is

$$(8_2) \quad q = v/|v|^2, \quad p = |v|^2 u - 2\tau v, \quad \text{where} \quad \tau = u \cdot v \quad (v \neq 0).$$

* This implies that the mapping of the two $2n$ -dimensional phase spaces (p, q) , (u, v) on each other is volume and orientation preserving ($\mu = +1$).

As far as the time derivatives are concerned, the applications (cf., e.g., §122–§124 bis; §498–501 bis) often warrant the combination of the transition from the configuration space $q = (q_1, \dots, q_n)$ to $v = (v_1, \dots, v_n)$ with the transition from t to another time variable, \bar{t} , which is defined by the condition that the local distortion of the time axis become proportional to the local distortion of the configuration space; so that

$$d\bar{t}/dt = dv_1 dv_2 \cdots dv_n / dq_1 dq_2 \cdots dq_n, \quad \text{i.e.,} \quad \bar{t}' = \det J, \quad (J = v_q).$$

As to an explicit rule for the introduction of \bar{t} , cf. §180. A particular case of $d\bar{t}/dt = \det J$ is the fundamental rule (11₂) §230, where $n = 2$.

† A transformation $s = f(r)$ is called involutory if its inverse is $r = f(s)$; so that $f(f(r)) = r$.

It is seen from (8₁)–(8₂) that

$$(8_0) \quad |q|^2 |v|^2 = 1, \quad |p|^2 |q|^2 = |u|^2 |v|^2; \quad p \cdot q + u \cdot v = 0 \\ \text{(i.e., } \sigma = -\tau \text{)}.$$

§51. If the degree of freedom is $n = 1$, one can write the completely canonical transformation (6) in the form

$$(9) \quad v = \int^u s(\bar{q}) d\bar{q}, \quad u = p/s(q), \quad \text{where } s = s(q) \neq 0$$

is a scalar function. A particular case of (9) is

$$(10) \quad v = sq, \quad u = p/s, \quad \text{where } s = \text{const.} \neq 0.$$

If the degree of freedom is $n = 2$, and the momenta $p_1, p_2; u_1, u_2$ and the coordinates $q_1, q_2; v_1, v_2$ are denoted by $\Xi, H; X, Y$ and $\xi, \eta; x, y$, respectively, the completely canonical transformation (6) reduces to

$$(11) \quad \begin{aligned} x &= x(\xi, \eta), & y &= y(\xi, \eta); \\ X &= \frac{y_\eta \Xi - y_\xi H}{x_\xi y_\eta - x_\eta y_\xi}, & Y &= \frac{-x_\eta \Xi + x_\xi H}{x_\xi y_\eta - x_\eta y_\xi}, \end{aligned}$$

where the denominator is $\det J (\neq 0)$. For instance, the canonical extension of the coordinate transformation which defines polar coordinates is

$$(12) \quad \begin{aligned} x &= \rho \cos \vartheta, & y &= \rho \sin \vartheta; \\ X &= P \cos \vartheta - \Theta \rho^{-1} \sin \vartheta, & Y &= P \sin \vartheta + \Theta \rho^{-1} \cos \vartheta, \end{aligned}$$

as seen by writing $\rho, \vartheta; P, \Theta$ for $\xi, \eta; \Xi, H$ in (11).

If one introduces the complex notations

$$(13) \quad z = x + iy, \quad \zeta = \xi + i\eta; \quad Z = X + iY, \quad \mathcal{Z} = \Xi + iH,$$

the coordinate transformation $x = x(\xi, \eta), y = y(\xi, \eta)$ appears as a mapping $z = z(\zeta)$ of two complex planes on each other. Suppose that $z = z(\zeta)$ is a regular analytic function.* Then the mapping is conformal everywhere, since $0 \neq \det J = |z_\zeta|^2$ by the Cauchy-Riemann equations $x_\xi = y_\eta, x_\eta = -y_\xi$. Thus, the completely canonical transformation (11) reduces, by (13), to

* This condition is not satisfied in (12), since $x + iy = \rho e^{i\vartheta}$ is not an analytic function of $\rho + i\vartheta$. However, one can choose $x + iy = e^{\xi + i\eta}$, put $e^\xi = \rho, \eta = \vartheta$, and then apply (9) to $s(q) = e^q$.

$$(14) \quad z = z(\zeta), \quad Z = Zz_\zeta(\zeta)/|z_\zeta(\zeta)|^2, \quad \text{where } z_\zeta \equiv dz/d\zeta \neq 0.$$

§52. Since $x + iy = z(\xi + i\eta)$, where $x_\xi = y_\eta$, $x_\eta = -y_\xi$, one has

$$(15_1) \quad x^2 + y^2 = |z(\xi + i\eta)|^2 \equiv |z|^2;$$

$$(15_2) \quad 4|z_\zeta|^2 = |z^2|_{\xi\xi} + |z^2|_{\eta\eta},$$

and, in view of (14) and (13),

$$(16_1) \quad X^2 + Y^2 = (\Xi^2 + H^2)/|z_\zeta|^2;$$

$$(16_2) \quad xY - yX = (|\frac{1}{2}z^2|_\xi H - |\frac{1}{2}z^2|_\eta \Xi)/|z_\zeta|^2;$$

finally, since $dz/dt \equiv z' = z_\zeta \zeta'$,

$$(17_1) \quad x'^2 + y'^2 = |z_\zeta(\xi + i\eta)|^2(\xi'^2 + \eta'^2);$$

$$(17_2) \quad xy' - yx' = |\frac{1}{2}z^2|_\xi \eta' - |\frac{1}{2}z^2|_\eta \xi'.$$

These formulae will now be applied to the Lagrangian function

$$(18) \quad L = \frac{1}{2}(x'^2 + y'^2) + (xy' - yx')f(x, y) + U(x, y),$$

where f , U are given functions (of class $C^{(2)}$) of $n = 2$ coordinates x, y . According to §15, the associated Hamiltonian function, H , is obtained by expressing $x'L_{x'} + y'L_{y'} - L$ in terms of $L_{x'}$, $L_{y'}$; x, y , instead of x', y' ; x, y . If X, Y denote the momenta $L_{x'}$, $L_{y'}$, then, by (18),

$$(19_1) \quad X = x' - yf, \quad Y = y' + xf; \quad (19_2) \quad x' = X + yf, \quad y' = Y - xf,$$

(19₂) being equivalent to the definition (19₁). Since $H = x'X + y'Y - L$, it is readily found from (19₂) and (18) that the Hamiltonian function is

$$(20) \quad H = \frac{1}{2}(X^2 + Y^2) - (xY - yX)f(x, y) - \{U(x, y) - \frac{1}{2}(x^2 + y^2)[f(x, y)]^2\}.$$

Introduce into (20) new coordinates ξ, η and momenta Ξ, H by means of (13)–(14), where $\zeta = \zeta(z)$ is the locally unique inverse of a given analytic function $z = z(\zeta)$. Since the transformation (14) is completely canonical, it transforms (20) into a Hamiltonian function K which is identical with (20) in virtue of (14); cf. §34. Hence, denoting K again by H , one sees from (16₁)–(16₂) that (20) is transformed by (14) into

$$(21) \quad H = |z_\zeta|^{-2} \left\{ \frac{1}{2}(\Xi^2 + H^2) - (|\frac{1}{2}z^2|_\xi H - |\frac{1}{2}z^2|_\eta \Xi)f - |z_\zeta|^2(U - \frac{1}{2}|z|^2 f^2) \right\},$$

where $|z_\zeta|^2$, $|z^2|_\xi$, $|z^2|_\eta$, $|z|^2$ and $f = f(x, y)$, $U = U(x, y)$ are thought of as expressed by means of $x + iy \equiv z = z(\xi + i\eta)$ as functions of ξ, η .

According to §10, the Lagrangian function \bar{L} into which the coordinate transformation $z = z(\zeta)$ transforms (18) is obtained by expressing L in terms of ξ, η ; so that $\bar{L} = L$ in virtue of $z = z(\zeta)$ and the derived relation $z' = z_\zeta(\zeta)\zeta'$. Hence, denoting \bar{L} again by L , one sees from (17₁)–(17₂) that (18) is transformed into

$$(22) \quad L = \frac{1}{2}|z_\zeta|^2 (\xi'^2 + \eta'^2) + (|\frac{1}{2}z^2|_\xi \eta' - |\frac{1}{2}z^2|_\eta \xi')f + U,$$

where $|z_\zeta|^2$, $|z^2|_\xi$, $|z^2|_\eta$ and $f = f(x, y)$, $U = U(x, y)$ are thought of as expressed by means of $x + iy \equiv z = z(\xi + i\eta)$ as functions of ξ, η .

§53. It is easily verified from the rules of §15 that (21) and (22) form an associated pair in the sense of §16, i.e., that (22) belongs to (21) in the same sense as (18) belongs to (20). Actually, this is clear for any extended canonical transformation, and for any n , from the last remark of §48.

If the degree of freedom is $n > 2$, then (8₁) is the only non-trivial analogue of (14), since it is known that, except for translations, rotations, reflections and changes of the unit of length, the inversion $v = q/|q|^2$ is the only conformal mapping of a Euclidean space of dimension $n > 2$ (Liouville).

In §54–§56, there will be collected for later use some classical coordinate transformations $v = v(q)$ of the type $z = z(\zeta)$; their canonical extensions then follow from (14) or (6).

§54. Let \mathbf{H}^{ξ_0} and \mathbf{E}^{η_0} denote the curves in the (x, y) -plane which correspond to the lines $\xi = \xi_0$ and $\eta = \eta_0$ of the (ξ, η) -plane if

$$(23) \quad x = -\mu + \xi^2 - \eta^2, \quad y = 2\xi\eta,$$

where μ is a given constant (not to be confused with a multiplier in §27). According to (13), one can write the coordinate transformation (23) in the form

$$(24) \quad \begin{aligned} x + iy \equiv z = z(\zeta) &\equiv -\mu + (\xi + i\eta)^2; \quad \text{so that} \\ |z_\zeta|^2 &= 4(\xi^2 + \eta^2). \end{aligned}$$

Thus, the condition $\det J \equiv |z_\zeta|^2 \neq 0$ of (14) is satisfied except at the point $\zeta = 0$, a point which belongs to $z = -\mu$ and represents, as does the point $\zeta = \infty$ which belongs to $z = \infty$, a branch point of

first order (i.e., one at which two sheets of the Riemann surface unite). Except for these branch points, the correspondence between the planes (x, y) and (ξ, η) is 1-to-2.

Correspondingly, (23) shows that if $\xi \neq 0$ then \mathbf{H}^ξ , and if $\eta \neq 0$ then \mathbf{E}^η , is a parabola such that $\mathbf{H}^\xi = \mathbf{H}^{-\xi}$ and $\mathbf{E}^\eta = \mathbf{E}^{-\eta}$; and that all these parabolas have the common focus $(x, y) = (-\mu, 0)$; finally, that their axes, when oriented from the focus towards the respective vertices, are the positively and negatively oriented x -axis, respectively; so that \mathbf{H}^0 and \mathbf{E}^0 are the (double) half-lines into which the x -axis is separated by the common focus. Hence, while the mapping (23)–(24) doubles the angles at $(\xi, \eta) = (0, 0)$, the curves \mathbf{H}^ξ and \mathbf{E}^η cross under right angles if $(\xi, \eta) \neq (0, 0)$; a fact which is clear from the conformity of the mapping also.

Thus, the coordinates ξ, η defined by (24) are the standard parabolic coordinates.

§55. The coordinate transformation (24), while rather simple locally, can lead to inconveniences in the large (cf. §451). A mapping which is locally equivalent to, but in the large often more convenient than, (24) results if one subjects $z + \mu$ and ζ in (24) to one and the same linear substitution, l ; choosing this l so as to transform $-\mu, 1 - \mu, \infty$ into $0, \infty, 1$, respectively, where μ is a given number. Thus, the transformation in question is

$$(25) \quad z = \frac{\zeta^2 + \mu(1 - \mu)}{2\zeta - (1 - 2\mu)}, \quad \text{since then} \quad l(z) = (l(\zeta))^2,$$

$$\text{where} \quad l(\zeta) = \frac{\zeta + \mu}{\zeta - 1 + \mu}.$$

According to (25), the correspondence between the planes (x, y) and (ξ, η) , where $z = x + iy$ and $\zeta = \xi + i\eta$, is again 1-to-2 except for two branch points P_1, P_2 of first order. Both of these belong, however, to finite z , and have image points Π_1, Π_2 which belong to finite ζ . In fact, from (25),

$$(26) \quad P_1: (-\mu, 0), \quad P_2: (1 - \mu, 0); \quad \Pi_1: (-\mu, 0), \quad \Pi_2: (1 - \mu, 0),$$

these Π_1, Π_2 being the points ζ of vanishing derivative z_ζ , and P_1, P_2 their z -images, i.e., the double points of the 1-to-2 mapping. Correspondingly, there is this time no branch point at infinity. In fact, (25) shows that the two distinct points

$$(27) \quad \Pi_0: (\xi, \eta) = (\tfrac{1}{2} - \mu, 0) \quad \text{and} \quad (\xi, \eta) = \infty \quad \text{belong to} \quad (x, y) = \infty.$$

Let $r_\nu = r_\nu(x, y)$ and $\rho_\kappa = \rho_\kappa(\xi, \eta)$, where $\nu = 1, 2$ and $\kappa = 0, 1, 2$, denote the distance between P_ν and a variable $P: (x, y)$, and between Π_κ and a variable $\Pi: (\xi, \eta)$ in the planes of z and ζ , respectively. Then r_1, r_2 and ρ_1, ρ_2 are bipolar coordinates in the respective planes, with P_1, P_2 and Π_1, Π_2 as poles. From (27) and (26),

$$(28) \quad \begin{aligned} \rho_0^2 &= (\xi + \mu - \tfrac{1}{2})^2 + \eta^2; & \rho_1^2 &= (\xi + \mu)^2 + \eta^2, \\ \rho_2^2 &= (\xi - 1 + \mu)^2 + \eta^2, \end{aligned}$$

while $r_1^2 = (x + \mu)^2 + y^2$, $r_2^2 = (x - 1 + \mu)^2 + y^2$. Hence, from (25),

$$(29_1) \quad r_1 = \tfrac{1}{2}\rho_1^2/\rho_0, \quad r_2 = \tfrac{1}{2}\rho_2^2/\rho_0; \quad (29_2) \quad |z_\zeta| = \tfrac{1}{2}\rho_1\rho_2/\rho_0^2.$$

§56. Another coordinate transformation which is again similar to, but more elaborate than, (23) is defined by

$$(30) \quad x = -\mu + \tfrac{1}{2} + \tfrac{1}{2} \cos \xi \cosh \eta, \quad y = \tfrac{1}{2} \sin \xi \sinh \eta,$$

where $\cosh w = \cos iw$, $\sinh w = -i \sin iw$. Thus, corresponding to (24) or (25),

$$(31) \quad x + iy \equiv z = z(\zeta) \equiv -\mu + \tfrac{1}{2} \{1 + \cos(\xi + i\eta)\};$$

($\mu = \text{const.} \geq 0$).

It is easily verified from (31) that

$$(32_1) \quad |z|^2 = (\tfrac{1}{2} - \mu)^2 + (\tfrac{1}{2} - \mu) \cosh \eta \cos \xi + \tfrac{1}{8}(\cosh 2\eta + \cos 2\xi);$$

$$(32_2) \quad |z_\zeta|^2 = \left| \sin \tfrac{1}{2}(\xi + i\eta) \cos \tfrac{1}{2}(\xi + i\eta) \right|^2 = \tfrac{1}{8}(\cosh 2\eta - \cos 2\xi).$$

In what follows, $(\xi, \eta) = \infty$ and $(x, y) = \infty$ will not be considered. This excludes, in particular, the logarithmical branch points of the Riemann surface of the inverse function $\zeta = \zeta(z)$. The remaining branch points, i.e., the (finite) ζ at which $z_\zeta \equiv -\tfrac{1}{2} \sin \zeta = 0$, belong to $\zeta = 0, \pm \pi, \pm 2\pi, \dots$ and are of the first order, since $z_{\zeta\zeta} \equiv -\tfrac{1}{2} \cos \zeta \neq 0$ at these ζ . Let S denote the (x, y) -plane, and Σ^k , where $k = 0, \pm 1, \pm 2, \dots$, the strip $2k\pi \leq \xi < 2k + 1\pi$ parallel to the η -axis in the (ξ, η) -plane, finally P_1, P_2 and Π_1^k, Π_2^k the pairs $(x, y) = (-\mu, 0)$, $(x, y) = (1 - \mu, 0)$ and $(\xi, \eta) = (2\pi k + \tfrac{1}{2}, 0)$, $(\xi, \eta) = (2\pi k, 0)$ of distinct points of S and Σ^k , respectively (so that P_1, P_2 are the same points as in §55). According to (31), the correspondence between S and Σ^k is 1-to-2 for every fixed k , save for the branch points P_1, P_2 and their images Π_1^k, Π_2^k ; the point Π_ν^k of Σ^k being mapped for every k and for $\nu = 1, 2$ on the single point P_ν of S .

In order to describe, as in §54, the curves $\xi = \text{const.}$ and $\eta = \text{Const.}$ in the (x, y) -plane, it is convenient to replace the essentially 1-to-2 correspondence (31) between S and Σ^k by an essentially 1-to-4 correspondence, as follows: Replace x, y by the bipolar coordinates r_1, r_2 which have P_1, P_2 as poles; so that, as at the end of §55,

$$(33) \quad r_1 = |(x + \mu)^2 + y^2|^{\frac{1}{2}} \geq 0, \quad r_2 = |(x - 1 + \mu)^2 + y^2|^{\frac{1}{2}} \geq 0.$$

Then $r_1 + r_2 \geq 1$, where $r_1 + r_2 = 1$, if and only if (x, y) lies between the poles P_1, P_2 on the x -axis; while $r_\nu = 0$ if and only if (x, y) is P_ν , where $\nu = 1, 2$. Excepting all points of the x -axis and only these points, the correspondence between (x, y) and (r_1, r_2) is 2-to-1, since the points $(x, y), (x, -y)$ and only these have the same bipolar coordinates (r_1, r_2) . This, when compared with the essentially 1-to-2 correspondence between S and a Σ^k , implies that there is an essentially 1-to-4 correspondence between (r_1, r_2) and the points (ξ, η) of every fixed strip Σ^k ; so that it is convenient to think of the strip Σ^k as consisting of four congruent half-strips.

Actually, the square roots (33) become uniformized* by ξ, η . In fact,

$$(34) \quad r_1 + r_2 = \cosh \eta, \quad r_1 - r_2 = \cos \xi;$$

so that r_1, r_2 are entire functions of ξ, η . For it is clear from (31) that $(x + \mu) \pm iy = \cos^2 \frac{1}{2}(\xi \pm i\eta)$, $(x - 1 + \mu) \pm iy = -\sin^2 \frac{1}{2}(\xi \pm i\eta)$; hence, (33) can be written as $r_1 = \cos \frac{1}{2}(\xi + i\eta) \cos \frac{1}{2}(\xi - i\eta)$, $r_2 = \sin \frac{1}{2}(\xi + i\eta) \sin \frac{1}{2}(\xi - i\eta)$, which proves (34).

Since all the strips Σ^k are equivalent, it is sufficient to consider the strip Σ^0 , i.e., the region $0 \leq \xi < 2\pi$, $-\infty < \eta < +\infty$ in the

* This holds for the coordinates ξ, η defined by (31) but not for those defined by (24) or by (25); as to (25), cf. (29₁) and (28).

It should be mentioned that (34) easily leads to the representation of $x'^2 + y'^2$ in terms of r'_1, r'_2 . First, it is seen from (32₂) and (34) that $|z'_\xi|^2 = r_1 r_2$, and so, from (17₁), that $x'^2 + y'^2 = r_1 r_2 (\xi'^2 + \eta'^2)$. On the other hand, it is clear from (34) that

$$\begin{aligned} r'_1 + r'_2 &= \eta' \sinh \eta, \quad r'_1 - r'_2 = -\xi' \sin \xi; \\ \sinh^2 \eta &= (r_1 + r_2)^2 - 1, \quad \sin^2 \xi = 1 - (r_1 - r_2)^2. \end{aligned}$$

Consequently, $x'^2 + y'^2 = r_1 r_2 (\xi'^2 + \eta'^2)$ reduces to

$$(35) \quad x'^2 + y'^2 = \frac{q_1^2 - q_2^2}{q_1^2 - \frac{1}{4}r_0^2} q_1'^2 + \frac{q_2^2 - q_1^2}{q_2^2 - \frac{1}{4}r_0^2} q_2'^2, \quad \text{where } \left. \begin{matrix} q_1 \\ q_2 \end{matrix} \right\} = \frac{r_1 \mp r_2}{2},$$

while r_0 denotes the fixed distance between the poles P_1, P_2 of the bipolar coordinates r_1, r_2 ; a distance which is $r_0 = 1$ in the present notation.

(ξ, η) -plane. For a given point (ξ_0, η_0) of Σ^0 , let \mathbf{H}^{ξ_0} and \mathbf{E}^{η_0} denote the curves in the (x, y) -plane which correspond, in virtue of (34) and (33), to the line $-\infty < \eta < +\infty$, $\xi = \xi_0$ and to the segment $0 \leq \xi < 2\pi$, $\eta = \eta_0$, respectively; so that the curves \mathbf{H}^ξ and \mathbf{E}^η are defined for $0 \leq \xi < 2\pi$ and $-\infty < \eta < +\infty$. Since r_1, r_2 are bipolar coordinates in the (x, y) -plane, with $P_1 = (-\mu, 0)$ and $P_2 = (1 - \mu, 0)$ as poles, it is clear from (34) that if η has a fixed non-vanishing value, \mathbf{E}^η is an ellipse with P_1 and P_2 as foci. Since $\cosh \eta$ is a steadily increasing function of $|\eta|$ and tends, as $\eta \rightarrow \pm 0$ and $\eta \rightarrow \pm \infty$, to $+1$ and $+\infty$, respectively, it is also clear from (34) that all ellipses \mathbf{E}^η together ($-\infty < \eta < +\infty$) cover the (x, y) -plane exactly twice, if one disregards the line segment \mathbf{E}^0 which joins P_1 with P_2 and connects the two families \mathbf{E}^η and $\mathbf{E}^{-\eta}$, where $\eta > 0$ and $\mathbf{E}^\eta = \mathbf{E}^{-\eta}$. (However, \mathbf{E}^η and $\mathbf{E}^{-\eta}$ have opposite orientations in virtue of their parameter representation (30) in terms of ξ). It is similarly seen, again from (34), that, unless $\xi = \frac{1}{2}\pi$ or $\xi = \frac{3}{2}\pi$, the curve \mathbf{H}^ξ is a branch of an hyperbola with P_1, P_2 as foci, and that all hyperbolic branches \mathbf{H}^ξ together ($0 \leq \xi < 2\pi$) cover the (x, y) -plane exactly twice, if one disregards $\mathbf{H}^{\frac{1}{2}\pi}$ and $\mathbf{H}^{\frac{3}{2}\pi}$ (lines which connect two families of hyperbolas or, rather, the four families of hyperbolic branches).

Thus, the mapping under consideration determines in the (x, y) -plane the so-called elliptic coordinates, defined in terms of confocal ellipses and hyperbolas. The parabolic case of §54 can be thought of as a limiting case.*

Canonical Matrices

§57. In what follows, an m -matrix will be thought of as consisting of m^2 constants.

If A is any m -matrix, the matrix series $\sum_{l=0}^{\infty} A^l/l!$, where $A^0 = (e_k^t)$, is convergent and defines an m -matrix which is denoted by e^A or $\exp A$. Clearly, $\exp(A') = (\exp A)'$ and, if T is non-singular, $\exp(TAT^{-1}) = T(\exp A)T^{-1}$. Furthermore, $e^{A+B} = e^A e^B$ whenever $AB = BA$. This implies, for $B = -A$, that $(e^A)^{-1}$ exists ($= e^{-A}$) for every A .

On choosing T so that TAT^{-1} becomes the Jordan normal form

* This may also be seen by writing $\cosh z$ in the form $\frac{1}{2}(Z + Z^{-1})$, where $Z = e^z$. In fact, the branch points of the inverse function of the rational function $\frac{1}{2}(Z + Z^{-1})$ are at $+1$ and -1 . If they were at a and b and one were to choose $a = 0$ and $b = \infty$, one would be led to the function Z^2 , which defines the parabolic coordinates.

of A , one sees from the definition of e^A that if α is a characteristic number of A , then e^α is a characteristic number of e^A and has the same multiplicity as α .

Unless the contrary is stated, all matrices are supposed to be real. On considering Jordan normal forms, one must, of course, leave the real field, if there are complex characteristic numbers.

§58. In the real field, the properties of symmetry, skew-symmetry and orthogonality are defined by $A' = A$, $A' = -A$ and $A' = A^{-1}$, respectively, where $A' = (a_i^k)$, if $A = (a_k^i)$. These properties are invariant under orthogonal transformations of A , and imply that all characteristic numbers of A are real, purely imaginary (inc. 0) and of absolute value 1, respectively. If $A' = A$ and if all characteristic numbers of A are positive (hence, $\det A > 0$), then A is called positive definite; while "non-negative definite" and "positive semi-definite" refer to an $A = A'$ with characteristic numbers which are all non-negative and all non-negative but not all positive, respectively. A matrix A is positive definite if and only if there exists a non-singular matrix B such that $A = BB'$; while $\det B = 0$ in the semi-definite case. If $A' = A^{-1}$ (hence, $\det A = \pm 1$), then A is called a rotation or a reflection according as $\det A = 1$ or $\det A = -1$.

The normal form of an arbitrary real m -matrix M under orthogonal transformations can be deduced from the fact that if $m > 2$, then there exists a rotation R such that, on placing $RM R^{-1} = (c_k^i)$, one has $c_k^1 = 0$ and $c_k^2 = 0$ for every $k > 2$.

§59. There exist for every non-singular m -matrix A exactly one positive definite P and exactly one orthogonal O such that $A = PO$ (where, $\det P$ being positive, $\det O = \pm 1$ is of the same sign as $\det A$).*

Since AA' is positive definite (§58), the existence and uniqueness of this "polar factorization" $A = PO$ follows immediately, if one shows that there exists for every given positive definite Q exactly one positive definite P such that $P^2 = Q$. For if $AA' = P^2$, where $P = P'$, the matrix O defined by $O = P^{-1}A$ is obviously such that $OO' = (e_k^i)$, and conversely. But orthogonal transformation of an

* If $m = 3$, the unique factorization PO of every non-singular A is familiar from the kinematics of continua, where it is shown that every linear deformation A of positive determinant can be decomposed into a unique rotation O and a unique dilatation P along three mutually perpendicular axes.

Similarly, if $m = 4$, the theorem implies the standard factorization of a Lorentz transformation of positive determinant into two three-dimensional Euclidean rotations and a positive definite binary Lorentz transformation.

arbitrary non-negative definite Q into a diagonal form shows that, whether Q does or does not possess multiple characteristic numbers, there exists exactly one non-negative definite P such that $P^2 = Q$. Since P is positive definite or semi-definite according as the same holds for P^2 , the proof is complete. (It may be mentioned that if $\det A = 0$, there exists exactly one positive semi-definite P but more than one orthogonal O such that $A = PO$.)

The factorization $A = PO$ is equivalent to a factorization $A = \mathbf{O}\mathbf{P}$, since $\mathbf{O} = O$, $\mathbf{P} = O^{-1}PO$. Clearly, $\mathbf{P} = P$ (and $\mathbf{O} = O$) if and only if $AA' = A'A$.

§60. Let C be a constant $2n$ -matrix ($m = 2n$). It will be called a canonical matrix if the conservative linear transformation $y = Cx$ is canonical in the sense of §27. Since the Jacobian matrix y_x is C , it is seen from §27 that C is a canonical matrix if and only if there exists a scalar multiplier $\mu \neq 0$ such that

$$(1_1) \quad CIC' = \mu I \quad (\mu \neq 0);$$

$$(1_2) \quad I = \begin{pmatrix} (0) & (e_k^i) \\ - (e_k^i) & (0) \end{pmatrix} = -I' = -I^{-1}.$$

This implies, by §32 and §31, that

$$(2_1) \quad \det C = \mu^n (\neq 0); \quad (2_2) \quad C'IC = \mu I; \quad (2_3) \quad C^{-1}IC^{-1'} = \mu^{-1}I;$$

so that C' and C^{-1} also are canonical matrices. In accordance with §34, a matrix C will be called completely canonical if (1₁) holds for $\mu = 1$ (which, by (2₁), implies that $\det C = 1$). For instance, I is, by (1₂), a completely canonical matrix.

On writing (2₃) in the form $\mu C^{-1} = IC'I^{-1}$, one sees that if α is a characteristic number of a completely canonical matrix C , then not only does the same hold for α^{-1} but α and α^{-1} have the same multiplicities, and even belong to invariant factors of the same degree. However, caution is necessary if $\alpha = \alpha^{-1}$, i.e., if $\alpha = \pm 1$. Thus, all that can be said is that the invariant factors of a completely canonical C which belong to an $\alpha \neq \pm 1$ occur in pairs corresponding to (α, α^{-1}) . The same holds, of course, also for pairs corresponding to $(\alpha, \bar{\alpha})$ if $\alpha \neq \bar{\alpha}$, i.e., if the (real) matrix C has a complex number α , hence also the complex conjugate $\bar{\alpha}$, as a characteristic number.

According to §31, the canonical matrices C form a group, and their multipliers μ are multiplicative on multiplication of the group elements.

§60 bis. The $2n$ -matrix $\exp(IH)$ is completely canonical for every symmetric $2n$ -matrix H . In fact, if $l = 0, 1, 2, \dots$ and $H' = H$, then, from (1₂), one has $[(IH)^l]' = (-HI)^l$, and so $[(IH)^l]' = I(-IH)^l I^{-1} \equiv [I(-IH)I^{-1}]^l$. This implies, by §57, that $[\exp(IH)]' = I[\exp(-IH)]I^{-1}$. Hence, it is clear from $\exp(-A) = (\exp A)^{-1}$ that (2₃) is satisfied by $C = \exp(IH)$, $\mu = 1$.

§61. It will be shown that if $C = PO$ is the unique polar factorization (§59) of a canonical matrix C of multiplier μ , then

$$(3) \quad OIO' = \operatorname{sgn} \mu \cdot I, \quad PIP' = |\mu| \cdot I, \quad \text{where } \operatorname{sgn} \mu = \mu/|\mu|.$$

In other words, P and O are again canonical and belong* to the multipliers $|\mu|$ and $\operatorname{sgn} \mu$.

In order to prove this, define, in terms of the data P, O, μ , four non-singular matrices $O_1, O_2; P_1, P_2$ by placing

$$(4) \quad O_1 = I, O_2 = \operatorname{sgn} \mu \cdot OIO^{-1}; \quad P_1 = P, P_2 = |\mu| \cdot O_2 P^{-1} O^{-1}.$$

Since $I' = I^{-1}$ by (1₂) and $O' = O^{-1}$ by assumption, while P , hence also P^{-1} , is positive definite, O_1 and O_2 are orthogonal, while P_1 and P_2 are positive definite. On the other hand, substitution of $C = PO$ into (2₂) gives $O^{-1} P I P O = \mu I$; a relation which, in view of the definitions (4) and (1₂), can be written in the form $P_1 O_1 = P_2 O_2$. It follows, therefore, from the uniqueness (§59) of the polar factorization of the non-singular matrix $P_1 O_1 = P_2 O_2$, that $O_1 = O_2$ and $P_1 = P_2$. But it is seen from (4) and (1₂) that $O_1 = O_2, P_1 = P_2$ can be written in the form (3).

The result, thus proved, can be interpreted as a one-to-one parametrization, $C = PO$, of the group of all canonical matrices C in terms of pairs P, O of canonical positive definite and canonical orthogonal $2n$ -matrices. Clearly, these O , but not these P , form a group.

Substituting in (1₁) an arbitrary orthogonal $2n$ -matrix $O = O^{-1}$ for C , and then using (1₂), one easily verifies that O is canonical if and only if

$$(5) \quad \begin{aligned} \text{either } O &= \begin{pmatrix} (a_k^i) & (b_k^i) \\ - (b_k^i) & (a_k^i) \end{pmatrix}, & \mu &= +1 \\ \text{or } O &= \begin{pmatrix} (a_k^i) & (b_k^i) \\ (b_k^i) & - (a_k^i) \end{pmatrix}, & \mu &= -1, \end{aligned}$$

* Choosing $C = P$, one sees that every positive definite canonical matrix is of positive multiplier.

where $(a_k^i), (b_k^i)$ are arbitrary n -matrices subject only to $\det O = \pm 1$.

§62. It is now easy to prove the fact announced in §32. The statement is that the obvious consequence $|\det C| = |\mu|^n$ of (1₁) may always be replaced by (2₁).

It is sufficient to prove this statement for all canonical $2n$ -matrices C whose multiplier is positive. The possibility of this reduction follows from (9₁)–(9₂), §31, if one multiplies any given canonical $2n$ -matrix C of negative multiplier by the matrix

$$G = \begin{pmatrix} (e_k^i) & (0) \\ (0) & - (e_k^i) \end{pmatrix}.$$

In fact, this G is easily verified to be a canonical $2n$ -matrix of the multiplier -1 and of determinant $(-1)^n$.

Accordingly, it is sufficient to prove (2₁) for every canonical C of positive multiplier. It follows, therefore, from §61 that it is sufficient to prove (2₁) for every positive definite $C = P$ and for every orthogonal $C = O$ of multiplier $+1$. But the determinant of a P is always positive; and the same holds, by the footnote to §61, for the multiplier of any $C = P$. Thus, all that remains to be shown is that a $C = O$ of multiplier $+1$ cannot have a negative determinant.

§62 bis. Since any $C = O$ of multiplier $+1$ has the form of the first of the two matrices (5), it is clear that if F denotes the (complex, unitary) $2n$ -matrix

$$F = \frac{1}{\sqrt{(2n)}} \begin{pmatrix} (e_k^i \sqrt{-1}) & (e_k^i) \\ (e_k^i) & (e_k^i \sqrt{-1}) \end{pmatrix}, \quad \text{then}$$

$$FOF^{-1} = \begin{pmatrix} (a_k^i + b_k^i \sqrt{-1}) & (0) \\ (0) & (a_k^i - b_k^i \sqrt{-1}) \end{pmatrix}.$$

Hence, $\det (FOF^{-1})$ is the product of the two complex conjugate numbers $\det (a_k^i \pm b_k^i \sqrt{-1})$, and so it cannot be negative. Since $\det (FOF^{-1}) = \det O$, the proof is complete.

§63. If a linear transformation $y = Cx$ of the $2n$ -vector $x = (x_i)$ into the $2n$ -vector $y = (y_i)$ is such as to transform the n -vectors $p = (p_i) \equiv (x_i)$ and $q = (q_i) \equiv (x_{i+n})$ of the momenta and coordinates into the respective n -vectors $u = (u_i) \equiv (y_i)$ and $v = (v_i) \equiv (y_{i+n})$, then C is completely canonical if and only if the trans-

formation of the coordinates is contragradient* to that of the momenta. This is clear from the last remark of §48 but follows more directly from §60. In fact, it is easily verified from (1₂) that (1₁) is satisfied by

$$(6) \quad C = \begin{pmatrix} (a_k^i) & (0) \\ (0) & (b_k^i) \end{pmatrix} \quad \text{if and only if} \quad (a_k^i)' = (b_k^i)^{-1}. \dagger$$

If the transformation matrix (a_k^i) of the momenta is identical with the transformation matrix (b_k^i) of the coordinates, (6) requires that $(a_k^i)' = (a_k^i)^{-1}$, which means that the n -matrix $(a_k^i) = (b_k^i)$ is orthogonal (cf. (15₁), §38).

§64. Let Q be a symmetric $2n$ -matrix of the particular form

$$(9) \quad Q = \begin{pmatrix} (r_k^i) & (0) \\ (0) & (s_k^i) \end{pmatrix}, \quad \text{where } Q' = Q, \text{ i.e., } r_k^i = r_i^k, \quad s_k^i = s_i^k.$$

Suppose further that at least one of the two symmetric n -matrices (r_k^i) , (s_k^i) , say (r_k^i) , is positive definite. Then there exists a completely canonical matrix C for which $C'QC$ becomes a diagonal matrix.

In order to prove this fact (which is fundamental in the theory of small vibrations), it is sufficient to show the existence of two matrices (a_k^i) , (b_k^i) which satisfy (6) and are such that both products

$$(10_1) \quad (a_k^i)'(r_k^i)(a_k^i);$$

$$(10_2) \quad (b_k^i)'(s_k^i)(b_k^i)$$

* Two linear transformations, determined by the matrices A and B , are called contragradient if $A = B'^{-1}$. In particular, the orthogonal matrices, and only these, determine linear transformations which are contragradient to themselves. Generally, one has to replace B by $A = B'^{-1}$ when passing from "point coordinates" to "line coordinates." Cf. also the pair of relations (31), §25.

† Examples of (6) are, for $2n = 6$,

$$(7) \quad (a_k^i) = \begin{pmatrix} 1 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix}, \quad (b_k^i) = \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & -1 \\ 0 & 0 & -1 \end{pmatrix},$$

and, if r_2, r_3 are arbitrary and $r_1 \neq 0$,

$$(8) \quad (a_k^i) = \frac{1}{r_1} \begin{pmatrix} 1 & -r_2 & -r_3 \\ 0 & r_1 & 0 \\ 0 & 0 & r_1 \end{pmatrix}, \quad (b_k^i) = \begin{pmatrix} r_1 & r_2 & r_3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The extension of (7) or (8) to $2n \geq 8$ is obvious.

are diagonal matrices. But, (r_k^t) being positive definite, §59 assures the existence of a non-singular symmetric (c_k^t) for which $(c_k^t)^2 = (r_k^t)$; in fact, (c_k^t) can be chosen as positive definite. Clearly, the product $(c_k^t)(s_k^t)(c_k^t)$ is a symmetric matrix and can, therefore, be represented in the form $(f_k^t)(d_k^t)(f_k^t)^{-1}$, where $(f_k^t)^{-1} = (f_k^t)'$ is an orthogonal and (d_k^t) a diagonal matrix. Thus, if (a_k^t) and (b_k^t) are defined by $(a_k^t) = (c_k^t)^{-1}(f_k^t)$ and $(b_k^t) = (c_k^t)(f_k^t)$, condition $(a_k^t)' = (b_k^t)^{-1}$ of (6) is satisfied, (10_2) becomes the diagonal matrix (d_k^t) , while (10_1) reduces to the unit matrix (e_k^t) , which is a diagonal matrix; so that the proof is complete.

§64 bis. Assume again that a given Q is of the form (9) but replace the additional assumption of §64 by the assumption that the matrices (r_k^t) , (s_k^t) are commutable. Then there exists again a completely canonical matrix C for which $C'QC$ becomes a diagonal matrix.

This criterion (which, in the particular case $(r_k^t) + (s_k^t) = (0)$, is fundamental in the theory of linear secular perturbation) may be proved by choosing C again in the particular form (6). In fact, the last remark of §63 shows that it is sufficient to prove the existence of an orthogonal (a_k^t) for which both matrices (10_1) , (10_2) become diagonal matrices if one chooses $(b_k^t) = (a_k^t)$. But the existence of such orthogonal (a_k^t) is known to be equivalent to the assumption that the symmetric matrices (r_k^t) , (s_k^t) are commutable.

Rotations

§65. For $m + m$ scalars a_i , b_i one has, if $\sum = \sum_1^m$, the obvious identity

$$(1) \quad \left| \begin{array}{cc} \sum a_i^2 & \sum a_i b_i \\ \sum a_i b_i & \sum b_i^2 \end{array} \right| = \frac{1}{2} \sum \sum \left| \begin{array}{cc} a_i & b_i \\ a_k & b_k \end{array} \right|^2; \quad \text{hence,}$$

$$(\sum a_i b_i)^2 \leq (\sum a_i^2)(\sum b_i^2).$$

In what follows, the vectors will be 3-vectors with reference to a Euclidean space, and it will be understood that, under the rotations of this space, the "vectors" transform as tensors, the rotations being represented by orthogonal 3-matrices, of determinant $+1$, which are formed by 3^2 constants.

Since $m = 3$, there is defined, not only the scalar product $a \cdot b = b \cdot a$, but also the vector product $a \times b = -b \times a$ of two vectors a , b . Placing $|c| = \sqrt{c^2} \geq 0$, where $c^2 = c \cdot c$, one has from (1)

$$(2) \quad \begin{aligned} |a \cdot b|^2 + |a \times b|^2 &= |a|^2 |b|^2; \text{ hence,} \\ |a \times b| &\leq |a| |b|, \quad |a \cdot b| \leq |a| |b|. \end{aligned}$$

The identity (2) may easily be generalized to the case of four 3-vectors:

$$(3) \quad (a \cdot c)(b \cdot d) - (a \cdot d)(b \cdot c) = (a \times b) \cdot (c \times d).$$

If $v = v(t)$ is a non-vanishing vector function of class $C^{(1)}$ on a t -interval, then the scalar $|v(t)|$ is of class $C^{(1)}$, since $|v|^2 = v^2$ implies that

$$(4) \quad |v| |v'| = vv' \quad (|v| \neq 0); \text{ hence, } ||v'| \leq |v'|, \text{ by (2).}$$

§66. Let an orthogonal 3-matrix Ω (of determinant $+1$) be given as a function $\Omega(t)$ of class $C^{(2)}$. Since $\Omega' \Omega$ is the unit matrix, $(\Omega' \Omega)' \equiv (0)$, and so $\Omega^{-1} \Omega' = -(\Omega^{-1} \Omega')'$. Accordingly, $\Omega^{-1} \Omega'$ is skew-symmetric, and so $\Omega = \Omega(t)$ determines a 3-vector $S = S(t)$ and a 3-matrix $\Sigma = \Sigma(t)$ for which

$$(5) \quad S = \begin{pmatrix} s_1 \\ s_2 \\ s_3 \end{pmatrix}, \quad \Sigma \equiv \Omega^{-1} \Omega' = \begin{pmatrix} 0 & -s_3 & s_2 \\ s_3 & 0 & -s_1 \\ -s_2 & s_1 & 0 \end{pmatrix} = -\Sigma';$$

$$\Omega' = \Omega^{-1}.$$

Thus, $\Sigma' = (\Omega' \Omega')' = \Omega' \Omega'' + \Omega'' \Omega' = \Omega^{-1} \Omega'' - \Sigma \Omega' \Omega' = \Omega^{-1} \Omega'' - \Sigma^2$, i.e.,

$$(6) \quad \begin{aligned} \Omega^{-1} \Omega'' &= \Sigma' + \Sigma^2, \text{ where* } \Sigma^2 = (s_i s_k - |S|^2 e_{ik}); \\ |S|^2 &= s_1^2 + s_2^2 + s_3^2. \end{aligned}$$

§67. Not only does every $\Omega(t) \equiv \Omega^{-1}(t)$ determine, by (5), a matrix $\Sigma(t) \equiv -\Sigma'(t)$, i.e., a vector $S(t)$, but one can also start with an arbitrary $S(t)$ and then determine an $\Omega(t)$ which satisfies (5); and this $\Omega(t)$ is uniquely determined by the given $S(t)$ and by an initial $\Omega(0)$ which can be chosen as an arbitrary orthogonal matrix (of determinant $+1$).

In fact, if $S(t)$, i.e., $\Sigma(t)$, is given, the requirement $\Omega^{-1} \Omega' = \Sigma$ of (5) represents for $\Omega(t)$ a homogeneous linear differential equation. Hence, there cannot exist more than one $\Omega(t)$ which belongs to $S(t)$ or $\Sigma(t)$ and reduces at $t = 0$ to a given matrix $\Omega(0)$. On the other hand, there always exists such an $\Omega(t)$, namely

* This representation of $\Sigma^2 = \Sigma \Sigma$, where $(e_{ik}) = \text{unit matrix}$, is clear from (5).

$$(7) \quad \Omega \equiv \Omega(t) = \Omega(0) \exp \int_0^t \Sigma(\bar{t}) d\bar{t}, \quad \text{since} \quad \Omega^{-1}\Omega' = \Sigma.$$

In fact, the integral of a skew-symmetric matrix $\Sigma(t)$ is again skew-symmetric, while an obvious modification of §60 bis shows that e^Θ is an orthogonal matrix of determinant +1 for every skew-symmetric Θ .

§68. In order to show that the vector $S = S(t)$ belonging to $\Omega = \Omega(t)$ is a vector in the sense of §65, one has merely to show [cf. (5)] that if the skew-symmetric matrix $\Omega'\Omega' = \Sigma \equiv \Sigma(t)$ belongs to $\Omega = \Omega(t)$ and, correspondingly, $\bar{\Omega}'\bar{\Omega}' = \bar{\Sigma} \equiv \Sigma(t)$ to $\bar{\Omega} = P\Omega P^{-1}$, where $P^{-1} = P' = \text{const.}$, then $\bar{\Sigma} = P\Sigma P^{-1}$. But $\bar{\Sigma} = \bar{\Omega}'\bar{\Omega}'$ can be written as

$$(P\Omega P')' (P\Omega P')' = (P\Omega' P') (P\Omega' P') = P\Omega'\Omega' P' = P\Sigma P^{-1}.$$

§69. That the relations of §66–§67 are covariant under the transformations $P = \text{const.}$ of the rotations group, will now become evident in itself, since it will be shown that the matrix operations of §65 are equivalent to operations with products of vectors.

To this end, let $\Xi = \Xi(t)$ denote the vector into which a vector $X = X(t)$ of the Euclidean space is transformed by the rotation $\Omega = \Omega(t)$ of this space; so that

$$(8) \quad \Xi = \Omega X; \quad \Omega'^{-1} = \Omega(t) = \Omega = \begin{pmatrix} o_{11} & o_{12} & o_{13} \\ o_{21} & o_{22} & o_{23} \\ o_{31} & o_{32} & o_{33} \end{pmatrix};$$

$$X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \quad \Xi = \begin{pmatrix} \xi \\ \eta \\ \zeta \end{pmatrix}.$$

Let $\Omega(t)$ and $X(t)$, hence also $\Xi(t)$, be of class $C^{(2)}$ in t .

Clearly, the signs of the components s_1, s_2, s_3 of $S = S(t)$ in the definition (5) of $\Sigma = \Sigma(t)$ are chosen so that

$$(9_1) \quad \Sigma X = S \times X; \quad (9_2) \quad \Sigma' X = S' \times X;$$

$$(9_3) \quad \Sigma^2 X = (S \cdot X)S - (S \cdot S)X,$$

where the cross and the dot refer to vector and scalar multiplications, respectively; while TX denotes, for $T = \Sigma, \Sigma', \Sigma^2$ (where $\Sigma' = d\Sigma/dt$, $\Sigma^2 = \Sigma\Sigma$), the vector into which the vector X is trans-

formed by the matrix T . Since differentiation of (8) gives $\Xi' = \Omega'X + \Omega X'$ and $\Xi'' = \Omega X'' + 2\Omega'X' + \Omega''X$, it is seen from (5) and (6) that

$$(10_1) \quad \Omega^{-1}\Xi' = X' + \Sigma X;$$

$$(10_2) \quad \Omega^{-1}\Xi'' = X'' + 2\Sigma X' + (\Sigma' + \Sigma^2)X.$$

Finally, from (8), (9₁) and (10₁),

$$(11_1) \quad \Omega^{-1}\Xi' = X' + S \times X;$$

$$(11_2) \quad \Omega^{-1}(\Xi \times \Xi') = X \times (X' + S \times X);$$

$$(11_3) \quad \Omega^{-1}\Xi'' = X'' + 2S \times X' + S' \times X + (S \cdot X)S - (S \cdot S)X$$

by (10₂), (9₃). It is understood that $A + B \times C + D$ denotes $A + (B \times C) + D$.

Notice that $S(t) \equiv 0$ holds, by (5) and (7), if and only if $\Omega(t) = \text{const.}$

§70. In what follows, the Euclidean space mentioned in §69 will be identified with the space of the vector Ξ occurring in (8); so that $X = \Omega^{-1}\Xi$ is the coordinate vector in the "rotating" coordinate system $X: (x, y, z)$ into which the orthogonal matrix $\Omega^{-1} = \Omega^{-1}(t)$ (of determinant $+1$) transforms the "non-rotating" coordinate system $\Xi: (\xi, \eta, \zeta)$. Correspondingly, $X = X(t)$ and $\Xi = \Xi(t) \equiv \Omega(t)X(t)$ can be thought of as given paths of one and the same particle in the two coordinate systems; so that the vectors Ξ' or Ξ'' and X' or X'' are, respectively, the absolute and relative velocities or accelerations of the particle.

Since the components of these velocity and acceleration vectors are parallel to the coordinate axes ξ, η, ζ and x, y, z of the non-rotating and rotating coordinate systems Ξ, X , respectively, it is clear from (8), i.e., from $X = \Omega^{-1}\Xi$, that the projections of the absolute velocity and of the absolute acceleration on the axes x, y, z of the rotating coordinate system are the components of the vector $\Omega^{-1}\Xi'$ and $\Omega^{-1}\Xi''$, respectively. This is the kinematical significance of (10₁), (10₂) or (11₁), (11₃).

§71. For a given path $\Xi = \Xi(t)$ of the particle in the non-rotating coordinate system $\Xi: (\xi, \eta, \zeta)$, one can always choose the rotation $\Omega(t)$ so that the particle is for every t in the (x, y) -plane of the rotating coordinate system $\Omega^{-1}(t)\Xi \equiv X: (x, y, z)$; i.e., so that $z(t) \equiv 0$. For this choice of $\Omega(t)$, the relations (11₁), (11₂) reduce, in view of (5) and (8), to

$$(12_1) \quad \Omega^{-1}\Xi' = \begin{pmatrix} x' - s_3 y' \\ y' + s_3 x' \\ s_1 y' - s_2 x' \end{pmatrix};$$

$$(12_2) \quad \Omega^{-1}(\Xi \times \Xi') = \begin{pmatrix} s_1 y'^2 - s_2 x' y' \\ s_2 x'^2 - s_1 x' y' \\ x' y' - y' x' + s_3(x'^2 + y'^2) \end{pmatrix},$$

since also $z'(t) \equiv 0$ if $z(t) \equiv 0$.

Notice that $z(t) \equiv 0$ can be satisfied for any given $\Xi = \Xi(t)$ by essentially different choices of $\Omega(t)$, since it is allowed to transform any given $\Omega(t)$ by an arbitrary $\Omega_0 = \Omega_0(t)$ which leaves the axis z of the rotating coordinate system $X: (x, y, z)$ unchanged.

§72. The condition that the (x, y) -plane of the rotating coordinate system (x, y, z) rotates within the (ξ, η) -plane of the non-rotating coordinate system (ξ, η, ζ) can be expressed in any of the three equivalent forms

$$(13_1) \quad \Omega = \begin{pmatrix} \cos \phi & -\sin \phi & 0 \\ \sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix};$$

$$(13_2) \quad \Sigma = \begin{pmatrix} 0 & -s_3 & 0 \\ s_3 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad s_3 = \phi'; \quad (13_3) \quad S = \begin{pmatrix} 0 \\ 0 \\ \phi' \end{pmatrix}.$$

In fact, (13₁) is an identity in t for a suitable $\phi = \phi(t)$ if and only if $z \equiv \zeta$. Furthermore, (13₂) is, in view of (7), necessary and sufficient for (13₁). Finally, (13₃) is, by (5), equivalent to (13₂).

§73. If the path $\Xi = \Xi(t)$ of the particle considered in §70 lies in a fixed plane of the non-rotating coordinate system $\Xi: (\xi, \eta, \zeta)$, one can choose this plane to be the (x, y) -plane of a rotating coordinate system $X: (x, y, z)$ which satisfies the requirement $z(t) \equiv 0$ of §71. Then (13₁) is satisfied, and so (13₃), (8) show that (11₁) and (11₃) reduce to

$$(14_1) \quad \Omega^{-1}\Xi' = \begin{pmatrix} x' - \phi' y' \\ y' + \phi' x' \\ 0 \end{pmatrix};$$

$$(14_2) \quad \Omega^{-1}\Xi'' = \begin{pmatrix} x'' - 2\phi'y' - \phi'^2x - \phi''y \\ y'' + 2\phi'x' - \phi'^2y + \phi''x \\ 0 \end{pmatrix},$$

while (11₂) reduces to the scalar relation

$$(14_3) \quad \xi\eta' - \eta\xi' = xy' - yx' + \phi'(x^2 + y^2).$$

In fact, $z(t) \equiv 0$, and so $z'(t) \equiv 0$, $z''(t) \equiv 0$.

§74. On comparing §72 with §68, one sees that a rotation defined by an $\Omega(t)$ is a rotation about a suitably chosen axis of invariable position if and only if there exists an orthogonal matrix P which is independent of t and such that all elements of the third row (and third column) of the skew-symmetric matrix $P\Sigma(t)P^{-1}$ vanish for every t , where $\Sigma = \Omega^{-1}\Omega'$.

§75. The last remark of §58 implies that every skew-symmetric 3-matrix can be transformed by an orthogonal matrix into a normal form in which all elements of the third row vanish. It follows, therefore, from §72 that in order that the rotation defined by $\Omega(t)$ be a rotation about some fixed axis, it is sufficient (but not necessary) that $\Sigma(t) = \text{const.}$, i.e., that all three components s_i of the vector S be independent of t . According to (7), the corresponding rotations $\Omega(t)$ are characterized by $\Omega(t) = \Omega(0)e^{t\Sigma}$, where Σ is an arbitrary skew-symmetric constant matrix.

§76. In what follows, the value of t will be thought of as arbitrarily fixed; so that the matrices occurring are considered as constants.

For an arbitrary skew-symmetric Θ , put

$$(15) \quad \Theta = \begin{pmatrix} 0 & -d_3 & d_2 \\ d_3 & 0 & -d_1 \\ -d_2 & d_1 & 0 \end{pmatrix}, \quad D = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix}; \quad \text{so that}$$

$$\Theta^2 = (d_i d_k) - |D|^2 E,$$

where E is the unit matrix and $|D| = (d_1^2 + d_2^2 + d_3^2)^{\frac{1}{2}} \geq 0$. Cf. (5)–(6).

It will be shown that a 3-matrix Ω is an orthogonal matrix of determinant $+1$ if and only if there exists a skew-symmetric matrix Θ such that*

* One can write (16₂) as $\Omega = E + \Theta \text{si } |D| + \frac{1}{2}\Theta^2 \text{si}^2 \frac{1}{2} |D|$, where $\text{si } \alpha = (\sin \alpha)/\alpha$.

$$(16_1) \quad \Omega = e^\Theta; \quad (16_2) \quad \Omega = E + \frac{\sin |D|}{|D|} \Theta + \frac{1 - \cos |D|}{|D|^2} \Theta^2.$$

First, it is easily verified from (15) that $\det(\lambda E - \Theta) = \lambda^3 + |D|^2 \lambda$. Since every matrix satisfies its characteristic equation, it follows that $\Theta^3 + |D|^2 \Theta = 0$; and so $\Theta^{n+3} = -|D|^2 \Theta^{n+1}$, where $n = 0, 1, \dots$. Consequently, $\Theta^{2n+1} = (-|D|^2)^n \Theta$, $\Theta^{2n+2} = (-|D|^2)^n \Theta^2$, and so

$$(17) \quad \begin{aligned} e^\Theta &\equiv \sum_{n=0}^{\infty} \frac{\Theta^n}{n!} \\ &= E + \frac{\Theta}{|D|} \sum_{n=0}^{\infty} \frac{(-1)^n |D|^{2n+1}}{(2n+1)!} + \frac{\Theta^2}{|D|^2} \left(1 - \sum_{n=0}^{\infty} \frac{(-1)^n |D|^{2n}}{(2n)!} \right). \end{aligned}$$

Since the last two series are those of $\sin |D|$ and $\cos |D|$, it follows that (16₂) is equivalent to (16₁).

Next, if Θ is the matrix (15) belonging to the particular values $d_1 = 0, d_2 = 0, d_3 = \phi$, the matrix e^Θ represented by (17) clearly reduces to (13₁). Hence, there exists for every Ω of the particular form (13₁) a $\Theta = -\Theta'$ which satisfies (16₁). If Ω is an orthogonal matrix of determinant $+1$ but not of the particular form (13₁), there exists, by the last remark of §58, an orthogonal P for which $P\Omega P^{-1}$ is of the particular form (13₁). But $\exp(P\Theta P^{-1}) = P e^\Theta P^{-1}$, by §57; furthermore, $P\Theta P^{-1}$ is skew-symmetric whenever P is orthogonal and Θ skew-symmetric. Consequently, there exists for every orthogonal Ω of determinant $+1$ a skew-symmetric Θ which satisfies (16₁). That the converse also holds, has already been observed at the end of §67.

§77. Let I_i denote, for $i = 1, 2, 3$, the matrix obtained from the general skew-symmetric matrix (15) by choosing $d_k = 1$ or $d_k = 0$ according as $i = k$ or $i \neq k$. Then an arbitrary Θ and an arbitrary Ω can be written as

$$(18_1) \quad \Theta = d_1 I_1 + d_2 I_2 + d_3 I_3; \quad (18_2) \quad \Omega = \exp(d_1 I_1 + d_2 I_2 + d_3 I_3),$$

by (16₁). Since (17) and (15) imply that

$$(19) \quad e^{\phi I_1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \phi & -\sin \phi \\ 0 & \sin \phi & \cos \phi \end{pmatrix}, \quad e^{\phi I_2} = \begin{pmatrix} \cos \phi & 0 & \sin \phi \\ 0 & 1 & 0 \\ -\sin \phi & 0 & \cos \phi \end{pmatrix},$$

$$e^{\phi I_3} = \begin{pmatrix} \cos \phi & -\sin \phi & 0 \\ \sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

the orthogonal matrix $e^{\phi I_i}$ represents the rotation of a Cartesian frame about its i -th coordinate axis by the angle ϕ or $-\phi$ according as $i = 1, 3$ or $i = 2$.

§78. It is clear from §57 that (18₂) is not the same thing as

$$(20) \quad \Omega = e^{\vartheta_1 I_1} e^{\vartheta_2 I_2} e^{\vartheta_3 I_3},$$

if $\vartheta_i = d_i$. It is, however, true that a 3-matrix Ω is orthogonal and of determinant $+1$ if and only if it can be represented by means of three numbers ϑ_i in the form (20). Actually, (20) is not essentially different from the standard (but unsymmetric) representation of Ω , given under (21) below.

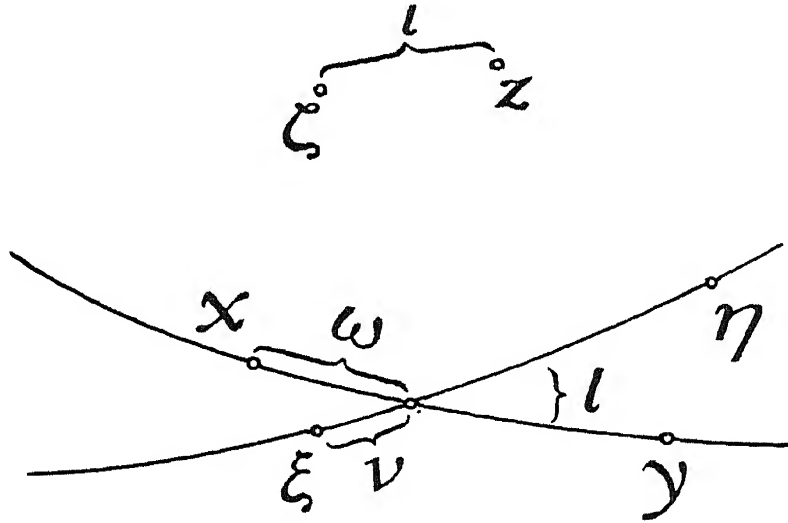


FIG. 1

It is clear from Fig. 1 that two arbitrarily given positions $E: (\xi, \eta, \zeta)$, $X: (x, y, z)$ of a Cartesian frame can be rotated into each other by rotating the frame first about its third coordinate axis by a suitable angle, then about the new position of the first axis by a suitable angle, finally about the resulting position of the third axis by a suitable angle. This means, in view of (19), that a 3-matrix Ω is orthogonal and of determinant $+1$ if and only if it can be represented in terms of three "Eulerian angles" ι, ν, ω as a matrix product

$$(21) \quad \Omega = e^{\nu I_3} e^{\iota I_1} e^{\omega I_3}.$$

Since $(e^{\Theta})^{-1} = e^{-\Theta}$ (cf. §57), it is seen from (15), (16₁) that the d_i of Ω^{-1} are the negatives of the d_i of Ω . On the other hand, $A^{-1}B^{-1}\Gamma^{-1} = (AB\Gamma)^{-1}$. Hence, (21) shows that

$$(22) \quad \Omega' = \Omega^{-1}: \{-\iota, -\omega, -\nu\}, \quad \text{if} \quad \Omega: \{\iota, \nu, \omega\}.$$

According to (19), the matrix product $e^{\nu I_3}e^{\iota I_1}$ is

$$(23) \quad e^{\nu I_3}e^{\iota I_1} = \begin{pmatrix} \cos \nu & -\cos \iota \sin \nu & \sin \iota \sin \nu \\ \sin \nu & \cos \iota \cos \nu & -\sin \iota \cos \nu \\ 0 & \sin \iota & \cos \iota \end{pmatrix}.$$

Multiplying (23) from the right by the matrix $e^{\omega I_3}$, one sees from (19) that the explicit representation of (21) is

$$(24) \quad \Omega = \begin{pmatrix} \cos \nu \cos \omega - \sin \nu \sin \omega \cos \iota & -\cos \nu \sin \omega - \sin \nu \cos \omega \cos \iota & \sin \nu \sin \iota \\ \sin \nu \cos \omega + \cos \nu \sin \omega \cos \iota & -\sin \nu \sin \omega + \cos \nu \cos \omega \cos \iota & -\cos \nu \sin \iota \\ \sin \omega \sin \iota & \cos \omega \sin \iota & \cos \iota \end{pmatrix}.$$

Clearly, (24) is equivalent to the fundamental formula of spherical trigonometry, the elements of (24) being the 3^2 direction cosines (cf. Fig. 1).

CHAPTER II

LOCAL AND NON-LOCAL QUESTIONS

Local notions	§ 79–§ 90
Hamiltonian and Lagrangian systems	§ 91–§102
Solutions and canonical transformations	§103–§118
Non-local notions	§119–§130
Points of stability	§131–§136
Characteristic exponents	§137–§154

Local Notions

§79. In what follows, \mathbf{X} will denote a given domain in the Euclidean space of an m -vector $x = (x_i)$, and $f(x)$ a given m -vector function $f = (f_i)$ which is, for some $\nu \geq 1$, of class $C^{(\nu)}$ on \mathbf{X} .

Denote by $|y|$ the Euclidean length of an m -vector y . It is clear that there exists for every point x^0 of \mathbf{X} a positive $\alpha = \alpha(x^0)$ which does not exceed b/B , where $b = b(x^0)$ is a positive number so small that the neighborhood $|x - x^0| < b$ of x^0 is contained in \mathbf{X} and $|f(x)|$ has for $|x - x^0| < b$ a finite least upper bound $B = B(x^0, b(x^0)) = B(x^0)$. It is also clear that $\alpha(> 0)$ can be chosen independent of x^0 , if x^0 is restricted to any fixed closed and bounded* subset of \mathbf{X} .

It is known that the system of m ordinary differential equations which is represented by

$$(1) \quad x' = f(x) \quad (' = d/dt)$$

has exactly one solution† path $x = x(t)$ which attains at an arbitrarily preassigned date $t = t^0$ an arbitrarily preassigned point x^0 of \mathbf{X} ; and that this solution $x = x(t)$ of (1) exists at least for $t^0 - \alpha < t < t^0 + \alpha$, where $\alpha = \alpha(x^0) = b/B$; finally, that $|x(t) - x^0| < b$ for $|t - t^0| < \alpha$.

If Δ denotes the differential operator

$$(2 \text{ bis}) \quad \Delta = \frac{\partial}{\partial t} + \sum_{i=1}^m f_i(x_1, \dots, x_m) \frac{\partial}{\partial x_i}, \text{ where } (f_i) = f, (x_i) = x,$$

it is seen from (1) that the solution paths $x = x(t)$ are characterized

* This means compactness, i.e., the applicability of the covering theorem of Heine-Borel.

† Notice that to a solution path corresponds, not only a locus in the x -space, but also a unique parametrization $x = x(t)$ of this locus.

among arbitrary paths $x = x(t)$ in the x -space by the fact that

$$(2) \quad (F(x; t))' = \Delta F(x; t) \text{ along solution paths } x = x(t),$$

where $F(x; t)$ is an arbitrary scalar or vector function which is of class $C^{(1)}$ in the $(m + 1)$ -dimensional $(x; t)$ -domain.

Since (1) does not contain t explicitly, it is clear from the uniqueness of the initial problem that a solution $x = x(t)$, when considered as a function of $x^0 = x(t^0)$, t^0 and t , is a function of x^0 and $t - t^0$ alone; say

$$(3) \quad x = x(x^0; t - t^0), \quad (x(x^0; 0) = x^0).$$

It is also known that, $f(x)$ being of class $C^{(\nu)}$ on \mathbf{X} , the m -vector function $x(x^0; t)$ belonging to (1) as well as the partial derivative $x_t(x^0; t)$ are of class $C^{(\nu)}$ on the $(m + 1)$ -dimensional $(x^0; t)$ -domain which is the product space† of \mathbf{X}^* and the t -interval $-\alpha < t < \alpha$, where \mathbf{X}^* is any domain formed by points x^0 of \mathbf{X} which have a bounded closure contained in \mathbf{X} and, correspondingly, $\alpha > 0$ is chosen independent of x^0 for all x^0 in \mathbf{X}^* .

Finally, it is known that, in the $(m + 1)$ -dimensional $(x^0; t)$ -domain under consideration, the Jacobian matrix x_{x^0} is non-singular, i.e.,

$$(4) \quad \det x_{x^0}(x^0; t) \neq 0 \quad (x_{x^0}(x^0; 0) = E = \text{unit matrix}).$$

Hence, on substituting (3) into (1), and differentiating the resulting n -vector identity in $(x^0; t - t^0)$ with respect to each of the n components of x^0 , one really sees from the differentiation rule of a determinant, that‡

$$(4 \text{ bis}) \quad (\log \det x_{x^0})' = \operatorname{div} f.$$

Notice that if $t - t^0$ is fixed and $|t - t^0| < \alpha$, then (4) assures that the mapping (3) of the x^0 -domain \mathbf{X}^* on an x -domain is of class $C^{[\nu]}$ in the sense of §5.

The inverse mapping is given by

$$(5) \quad x^0 = x(x; t^0 - t),$$

where x is the same function sign as in (3). In fact, there belongs to every point of \mathbf{X} at every date t exactly one solution path; so that

† Cf. the footnote to §9.

‡ The divergence, $\operatorname{div} f$, of $f = f(x)$ is defined as the trace of the Jacobian matrix f_x (as to the trace of a matrix, cf. the footnote to §137).

the equivalence of (5) and (3) follows by interchanging the initial and final states.

Needless to say, the transition from (3) to (5) is legitimate only when x^0 is restricted to some domain \mathbf{X}^* possessing a bounded closure contained in \mathbf{X} , while $|t - t^0|$ is supposed to be less than a constant depending on \mathbf{X}^* . In particular, one cannot be sure that there exists a fixed $t (\neq t^0)$ such that the function (3) is defined at this t and for every x^0 contained in \mathbf{X} . This situation leads to obvious complications; complications which will not always be emphasized but must not be forgotten.

§80. Let $G = G(x; t)$ be an l -vector function of class $C^{(1)}$ on the $(m + 1)$ -dimensional $(x; t)$ -domain under consideration, and suppose, (i): that there exists in this domain at least one point $(x; t)$ at which $G(x; t) = 0$, and, (ii): that if $x = x(t)$ is any given solution path of (1), then $G(x(t); t) = 0$ either holds for every t or for no t along the solution path. Then the system of l relations which is represented by $G(x; t) = 0$ is called an invariant system of (1). It is clear from (2) that (i)–(ii) can also be expressed by requiring, (i bis): that the l -vector condition $G(x; t) = 0$ is not contradictory within the $(m + 1)$ -dimensional $(x; t)$ -domain and, (ii bis): that $\Delta G(x; t) = 0$ becomes an identity in $(x; t)$ in virtue of $G(x; t) = 0$.

A scalar invariant system (where $l = 1$) is called an invariant relation. If $l > 1$, the l scalar relations which constitute the invariant system $G(x; t) = 0$ need not be invariant relations. Examples to this effect are implied by the remark that if $x = \xi(t)$ is any particular solution path of (1), then $G(x; t) = 0$, where $G(x; t) \equiv x - \xi(t)$, obviously is an invariant system of $l = m$ equations.

§81. A set \mathbf{X}^* of points x which is contained in the x -domain \mathbf{X} and contains at least one point is called an invariant set of (1) if it has the following property: There exists for every point x^* of \mathbf{X}^* a sufficiently small positive $\rho = \rho(x^*)$ in such a way that if $x = x(t)$ is any solution path for which $x^* = x(t^*)$ holds for a suitable $t = t^*$, then the point $x(t)$ is a point of \mathbf{X}^* for all those t for which $|x(t) - x^*| < \rho$.

It is clear that if an invariant relation $G = 0$ is conservative in the sense of §18, i.e., such that G is a function of x alone (instead of being a function of x and t), then $G(x) = 0$ is the equation of an invariant set. Actually, the notions of an invariant set and of a conservative invariant system seem to be hardly different. How-

ever, a *closed* invariant set $\mathbf{X}^+ : G(x) = 0$ of (1) can have a rather complicated structure, even if the functions $f(x)$ and $G(x)$ are very smooth (so that the question becomes of interest only under the restriction of analyticity).

§82. According to §80, a relation $G(x; t) = 0$ belonging to a scalar function $G \not\equiv 0$ of class $C^{(1)}$ is an invariant relation if and only if $G(x; t) = 0$ determines in the $(x; t)$ -domain an hypersurface on which the function $\Delta G(x; t)$ of $(x; t)$ vanishes identically. It is possible that a scalar function $F(x; t)$ of class $C^{(1)}$ is such that the function $\Delta F(x; t)$ of $(x; t)$ vanishes not only on the hypersurface $F(x; t) = 0$ but on the whole $(x; t)$ -domain. In contrast with the situation in §80, this will be the case if and only if $\Delta F(x; t) = 0$ is an identity in $(x; t)$ not merely in virtue of $F(x; t) = 0$ but in itself; so that $F(x; t)$ is a solution of the linear partial differential equation $\Delta F = 0$ defined by (2 bis). Since every initial condition $(x^0; t^0)$ determines a solution path $x = x(t)$, it is clear from (2) that this will be the case if and only if $(F(x(t); t))' = 0$, i.e., $F(x(t); t) = c = \text{const.}$, holds along every fixed solution path $x = x(t)$ of (1). One then calls the scalar function $F(x; t)$ or the relation $F(x; t) = c$, where the constant c is unspecified, an integral of (1), provided that the function $F(x; t)$ is not a constant on the $(m + 1)$ -dimensional $(x; t)$ -domain.

It is understood that the value of c which belongs to any given solution path $x = x(t)$ is, in view of $F(x^0; t^0) = c$, a function of the initial conditions $x^0 = x(t^0)$, t^0 . If c has a fixed value c_0 , then $F(x; t) = c_0$ is not an integral but, when written in the form $G(x; t) \equiv F(x; t) - c_0 = 0$, merely an invariant relation. In fact, if the function $G(x; t)$ is an integral, it must not contain an integration constant.

In accordance with §18, one calls an integral $F(x; t)$ conservative if it does not contain t . Then $F(x) = c_0$, where $c_0 = F(x^0)$, is called an integral hypersurface through $x = x^0$. This "hypersurface" can consist of the single point $x = x^0$ and is always an invariant set (§81).

It is obvious that any scalar function of integrals of (1) is again an integral, provided that the function is of class $C^{(1)}$ and does not become independent of $(x; t)$. Consequently, one can define l integrals F_1, \dots, F_l to be independent if the functions $F_1(x; t), \dots, F_l(x; t)$ are independent in the local sense of §18.

It is clear from the inversion (5) of (3), that the scalar functions which constitute the components of the m -vector function $\mathbf{x}(x; t^0 - t)$

represent m integrals of (1), and that these m integrals are, in view of (4), independent integrals.

The m independent integrals just mentioned cannot be all independent of t , unless $f(x) \equiv 0$. For suppose that (1) has m independent conservative integrals, say $F_1(x) = c_1, \dots, F_m(x) = c_m$. Then every solution path must be an intersection of m hypersurfaces $F_i(x) = c_i$, where $c_i = F_i(x(t^0))$. Since the m functions $F_i(x)$ are independent and the hypersurfaces lie in the m -dimensional x -space, it follows, by placing $c = (c_i)$, that $x(t) = c$ along every solution path $x = x(t)$. This means, in view of (1), that $f(x) \equiv 0$ in the x -space. Conversely, if $f(x) \equiv 0$, then $x_1 = c_1, \dots, x_m = c_m$ represent m independent conservative integrals of (1).

While there exist m conservative independent integrals only when $f(x) \equiv 0$, there always exist $m - 1$ conservative independent integrals $F_1(x), \dots, F_{m-1}(x)$. In order to see this, it is sufficient to eliminate $t_0 - t$ (in a suitable manner) between the m independent integrals which constitute the components of the m -vector relation (5). It is understood that the resulting $m - 1$ independent integrals $F_1(x), \dots, F_{m-1}(x)$ have a purely local significance not only with regard to t (cf. the end of §79) but, in view of the elimination process, with regard to x also.

§83. Notwithstanding the complications pointed out at the end of §79, one speaks sometimes of the manifold of all solutions $x = x(t)$ of (1) and calls, correspondingly, (3) the general solution (on the other hand, (5) represents m integrals, if t^0 is thought of as fixed).

A solution $x = x(t)$ of (1) is called an equilibrium solution of (1) if the path $x = x(t)$ in the x -space is represented by a single point; namely, by the point $x = x^0$, where $x^0 = x(t^0)$. This will be the case if and only if $f(x^0) = 0$. In fact, $x'(t) = 0$ cannot hold for a single $t = t^0$ unless it holds for every t , i.e., unless $x(t) \equiv x^0$. For if $x = x(t)$ satisfies (1) on a t -interval containing $t = t^0$, and if $x'(t^0) = 0$, then $0 = f(x^0)$; and so $x(t) \equiv x^0$ is one, hence the only, solution of (1) which satisfies the initial condition $x(t^0) = x^0$. Correspondingly, a point $x = x^0$ of the x -space is called an equilibrium point if $f(x^0) = 0$. In particular, the exceptional case of m independent conservative integrals (§82) is the case in which every point x is an equilibrium point.

It should be mentioned that if a solution $x = x(t)$ of (1) is not an equilibrium solution, the corresponding solution path in the x -space has at every t a tangent and is free of cusps. For if this did not hold

for some $t = t^0$, one would have $x'(t^0) = 0$, hence $x'(t) \equiv 0$, i.e., $x(t) \equiv x(t^0)$.

§84. Without loss of generality, choose $t^0 = 0$; so that the general solution (3) of (1) appears as $x = x(x^0; t)$, where $x^0 = x(x^0; 0)$. Suppose that a given particular solution $x = x(\bar{x}^0; t)$ which belongs to a fixed $x^0 = \bar{x}^0$ is known to exist not only on the small t -interval supplied by the local existence theorem (§79) but on a larger t -interval, say $0 \leq t \leq M$, where it is understood that the point $x = x(\bar{x}^0; t)$ is, for $0 \leq t \leq M$, a point of the x -domain \mathbf{X} introduced in §79. It will be shown that, no matter how large is the given number $M (\neq \infty)$, one can choose a $\delta > 0$ so small that all those solutions $x = x(x^0; t)$ of (1) exist on $0 \leq t \leq M$ which belong to any initial condition $x(x^0; 0) = x^0$ satisfying the inequality $|x^0 - \bar{x}^0| < \delta$.

To this end, let \mathbf{X}_η denote, for an arbitrary $\eta > 0$, the domain of those points x of the x -space for which $|x - x(\bar{x}^0; t)| < \eta$ holds for at least one t satisfying $0 \leq t \leq M$. Since the set \mathbf{X} introduced in §79 is open, one can choose $\eta > 0$ so small that \mathbf{X}_η is contained in a closed and bounded subset of \mathbf{X} . Then the positive number α of §79 can be so chosen as to be valid for every point $x = x_0$ of \mathbf{X}_η . In other words, if $x = x_0$ is any point of \mathbf{X}_η , and $t = t_0$ any point of the t -axis, the solution $x = x(t)$ with the initial condition $x(t_0) = x_0$ exists at least for $t_0 - \alpha < t < t_0 + \alpha$, where α is independent of x_0 and t_0 . Thus, on choosing t_0 within $0 \leq t \leq M$, and noting that x_0 and t_0 together determine exactly one local solution, one sees that the balance of the proof follows from the covering theorem of Heine-Borel. Since $x(x^0; t)$ is, by §79, a continuous function, hence uniformly continuous on every closed and bounded set, there follows for every $\epsilon > 0$ the existence of a $\delta = \delta_\epsilon > 0$ such that $|x(x^0; t) - x(\bar{x}^0; t)| < \epsilon$ for $0 \leq t \leq M$ whenever $|x^0 - \bar{x}^0| < \delta$.

Notice that this holds no matter how long is the fixed finite t -interval $0 \leq t \leq M$ on which the given particular solution $x(\bar{x}^0; t)$ is supposed to exist.

§85. Since $x(x^0; t)$ is, by §79, of class $C^{(\nu)}$, where $\nu \geq 1$, it is clear from Taylor's formula that*

$$(6) \quad x(x^0; t) = x(\bar{x}^0; t) + R(t)(x^0 - \bar{x}^0) + o(|x^0 - \bar{x}^0|)$$

holds uniformly for $0 \leq t \leq M$ as $x^0 \rightarrow \bar{x}^0$, where $R(t)$ denotes the Jacobian matrix of $x(x^0; t)$ with respect to x^0 at $x^0 = \bar{x}^0$, i.e.,

* As to the symbol o , cf. the footnote to §11.

$$(7) \quad R(t) = (x_{x^0}(x^0; t))_{x^0=\bar{x}^0}; \quad \text{so that} \quad R(0) = E, \quad \text{by (4).}$$

One can interpret (6) as an approximate representation of the general solution $x(x^0; t)$.

It is of fundamental significance that, without knowing the general solution of the system (1), one can determine the approximative representation (6), i.e., the matrix (7), by knowing the general solution of a *linear* system

$$(8) \quad \xi' = A(t)\xi,$$

where $A(t)$ is a known m -matrix function of t , namely, the Jacobian matrix of $f(x)$ with respect to x along the given particular solution $x = x(\bar{x}^0; t)$ of (1):

$$(9) \quad A(t) = (f_x(x))_{x=x(\bar{x}^0; t)}.$$

In fact, let ξ_i be the i -th component of an m -vector $\xi = \xi(t)$ which satisfies (8), and denote, for a fixed k ($= 1, \dots, m$), by $\xi^k(t)$ that particular solution of (8) which satisfies the m initial conditions $\xi_i^k(0) = e_{ik}$, where $i = 1, \dots, m$ and (e_{ik}) is the unit matrix E . Now, if one knows these m solutions $\xi^k(t)$ of (8), one also knows the matrix $R(t)$ occurring in (6), since the vector $\xi^k(t)$ is the k -th column of $R(t)$. This statement is equivalent to

$$(10) \quad R'(t) = A(t)R(t), \quad \text{since} \quad R(0) = E,$$

by (7). And the truth of (10) may be proved as follows:

According to §79, not only $x(x^0; t)$ but also $x_t(x^0; t) \equiv x'(x^0; t)$ is of class $C^{(\nu)}$, where $\nu \geq 1$. Hence, corresponding to (6),

$$(11) \quad x'(x^0; t) = x'(\bar{x}^0; t) + R'(t)(x^0 - \bar{x}^0) + o(|x^0 - \bar{x}^0|)$$

holds uniformly for $0 \leq t \leq M$, as $x^0 \rightarrow \bar{x}^0$. On the other hand, $x(t) = x(x^0; t)$ is a solution of (1) for an arbitrary x^0 and for the particular $x^0 = \bar{x}^0$; so that, by subtraction,

$$x'(x^0; t) - x'(\bar{x}^0; t) = f(x(x^0; t)) - f(x(\bar{x}^0; t)).$$

But from (6), from the definition (9), and from Taylor's formula,

$$f(x(x^0; t)) = f(x(\bar{x}^0; t)) + A(t)R(t)(x^0 - \bar{x}^0) + o(|x^0 - \bar{x}^0|).$$

Hence, $x'(x^0; t) - x'(\bar{x}^0; t) = A(t)R(t)(x^0 - \bar{x}^0) + o(|x^0 - \bar{x}^0|)$, or, by (11),

$$R'(t)(x^0 - \bar{x}^0) + o(|x^0 - \bar{x}^0|) = A(t)R(t)(x^0 - \bar{x}^0) + o(|x^0 - \bar{x}^0|).$$

Since x^0 is an arbitrary constant vector close to \bar{x}^0 , the proof of (10) is complete.

§86. According to (9), the coefficient matrix $A(t)$ of the system (8) of m homogeneous linear scalar differential equations for $\xi = (\xi_i)$ is, for a given system (1), uniquely determined by the given solution $x = x(\bar{x}^0; t)$ alone. This particular solution of (1) will from now on simply be denoted by $x = \bar{x}(t)$. The system (8) with the coefficient matrix (9) is called the "system of Jacobi equations (or equations of variation) associated with the given solution $x = \bar{x}(t)$ of (1)," while any solution $\xi = \xi(t)$ of (8), and not only one of the m solutions $\xi = \xi^k(t)$ considered in §85, is called "a displacement of the solution $x = \bar{x}(t)$ of (1) with reference to (1)." What is actually meant is an infinitesimal displacement, since the terminology quoted is intended merely to describe the following fact:

Let $\xi = \xi(t)$ be any m -vector function of class $C^{(1)}$ on an interval $0 \leq t \leq M$, and let $\epsilon > 0$ be a small parameter independent of t . Then the function $\bar{x}(t) + \epsilon\xi(t)$ of t satisfies (1) with an error of an order higher than the order of ϵ if and only if $\xi(t)$ is a displacement of the solution $x = \bar{x}(t)$ of (1). In other words, a given m -vector $\xi(t)$ will or will not have the property that

$$(12) \quad (\bar{x}(t) + \epsilon\xi(t))' = f(\bar{x}(t) + \epsilon\xi(t)) + o(\epsilon), \quad \epsilon \rightarrow 0,$$

holds uniformly for $0 \leq t \leq M$ according as $\xi(t)$ is or is not a solution of (8). In order to prove this, it is sufficient to observe that, by (9) and by Taylor's formula, $f(\bar{x}(t) + \epsilon\xi(t)) = f(\bar{x}(t)) + \epsilon A(t)\xi(t) + o(\epsilon)$; so that, since $\bar{x}'(t) = f(\bar{x}(t))$ by assumption, (12) is equivalent to $\epsilon\xi'(t) = \epsilon A(t)\xi(t) + o(\epsilon)$. Since $\xi(t)$ and $A(t)$ do not depend on ϵ , it follows that (12) is equivalent to (8).

§87. Let $x = x(t; \epsilon)$ be an m -vector function of class $C^{(1)}$ on a rectangle $0 \leq t \leq M$, $0 \leq \epsilon \leq \text{const.}$, and suppose that $x(t; \epsilon)$ is a particular solution of (1) for every fixed ϵ and reduces at $\epsilon = 0$ to the solution $\bar{x}(t)$ to which (8), (9) belong. Then the partial derivative

$$(13) \quad \xi(t) = x_\epsilon(t; 0), \quad (x(t; \epsilon) = \bar{x}(t) \text{ for } \epsilon = 0),$$

is a solution of (8). The proof is the same as at the end of §86.

Since every solution $\bar{x}(t)$ of (1) can, by §84, be embedded into suitably chosen families $x(t; \epsilon)$ of solutions, and since, in particular, the m solutions $\xi^k(t)$ of (8) which were considered in §85 are of the type (13), one readily sees that every solution $\xi(t)$ of (8) can be represented by means of families $x(t; \epsilon)$ of solutions of (1) in the form (13).

If $F(x)$ is an integral of (1), then $F(x(t; \epsilon))$ is, by §82, a function of ϵ alone. Hence, on differentiating $F(x(t; \epsilon))$ with respect to ϵ at $\epsilon = 0$, one sees from (9) that the scalar product

$$(14) \quad \xi(t) \cdot F_x(\bar{x}(t)) = \text{const.}$$

along the solution (13) of (8), and so along any given solution $\xi(t)$ of (8).

Since $\bar{x}(t + \text{const.})$ is, for every solution $\bar{x}(t)$ of (1), again a solution, application of (13) to the family $x(t; \epsilon) = \bar{x}(t + \epsilon)$ shows that (8) always admits the solution

$$(15) \quad \xi = \bar{x}'(t).$$

§88. If $y = y(x)$ is a mapping of class $C^{[2]}$ of the x -domain on a y -domain (cf. §5), the system (1) and its solution path $x = \bar{x}(t)$ are transformed into a system $y' = g(y)$ and a corresponding solution path $y = \bar{y}(t)$. Let $\eta(t)$ denote an arbitrary displacement of $\bar{y}(t)$ with reference to $y' = g(y)$; so that, corresponding to (8), (9),

$$(16) \quad \eta' = B(t)\eta;$$

$$(17) \quad B(t) = (g_y(y))_{y=\bar{y}(t)}.$$

Thus, if $S(t)$ denotes the matrix which belongs to $B(t)$ in the same way as (7) does to $A(t)$, then $S'(t) = B(t)S(t)$, by (10). The explicit connection between $R(t)$ and $S(t)$ is quite involved and cannot be expressed in terms of the Jacobian matrix $y_x \equiv J = J(t)$ of the mapping $y = \bar{y}(x)$ along the solution path $x = \bar{x}(t)$. Correspondingly, the connection between the coefficient matrices (9), (17) of the respective Jacobi systems (8), (16) is not expressible in terms of J alone.

Fortunately, the matrix differential equation $T'(t) = B(t)T(t)$ has a solution $T(t)$ which is more easily obtainable than the particular solution $T(t) = S(t)$, found above, and can be used for the same purposes. In fact, $J(t)R(t)$ also is a $T(t)$, i.e.,

$$(18) \quad \begin{aligned} \tilde{R}'(t) &= B(t)\tilde{R}(t) \quad \text{holds for} \\ \tilde{R}(t) &= J(t)R(t), \quad \text{where } J(t) = y_x(\bar{x}(t)); \quad \det J \neq 0. \end{aligned}$$

This is easily verified from (9), (10), (17) and from the representation of $g(y)$ in terms of $f(x)$ and of the Jacobian matrix $J = y_x$.

§89. In order that the Jacobi system (8) belonging to $x = \bar{x}(t)$ has a constant coefficient matrix A , it is, by (9), sufficient that the solution $\bar{x}(t)$ of (1) be independent of t , i.e., be an equilibrium solution.

In this case, the integration of (8) depends merely on the determination of the characteristic numbers and invariant factors of A .

The characteristic numbers of A , i.e., the roots s of $\det(sE - A) = 0$, are called the characteristic exponents of $\xi' = A\xi$. An s is said to be of stable type if it is purely imaginary (incl. 0). Clearly, every characteristic exponent must be of stable type if every solution $\xi(t)$ of $\xi' = A\xi$ remains bounded as $t \rightarrow \pm \infty$. The converse is not true, since in case of at least one multiple invariant factor the general solution of $\xi' = A\xi$ is not free of "secular" terms.

Clearly, $\xi' = A\xi$ is identical with its Jacobi system with regard to any of its solutions $\xi = \bar{\xi}(t)$.

§90. It has been assumed since §79 that (1) does not contain t explicitly. This is not a loss of generality, provided that one considers t as an $(m + 1)$ -st x_i . For if instead of (1) one has to deal with $x' = f(x; t)$, where $f = (f_i)$, $x = (x_i)$ and $i = 1, \dots, m$, then, on placing $f_0 \equiv 1$ and $x_0 \equiv t$ (so that $x_0^0 = t^0$), one can replace $x' = f(x; t)$ by $*x' = *f(*x)$, where $*f = (f_j)$, $*x = (x_j)$ and $j = 0, 1, \dots, m$.

For instance, one can say that (14) is an integral of (8) unless $F_x(\bar{x}(t)) = 0$ for every t ; cf. §82.

Hamiltonian and Lagrangian Systems

§91. If $H = H(x; t)$ is a Hamiltonian function for which $H_x(x; t)$ is of class $C^{(1)}$ in the $(2n + 1)$ -dimensional $(x; t)$ -domain, then, if use is made of the notations of §19, the system

$$(1) \quad x' + 1H_x(x; t) = 0, \text{ i.e., } p' = -H_q(p, q; t), \quad q' = H_p(p, q; t) \\ (I^{-1} = -I),$$

is called the corresponding system of Hamiltonian equations.

It is clear that two such systems are identical if and only if the difference of the two $H(x; t)$ is independent of x . Correspondingly, if (1) is conservative, i.e., if $H_x(x; t)$ is independent of t , one can assume that $H(x; t)$ is conservative, i.e., that $H_t \equiv 0$.

Placing $f(x; t) = -1H_x(x; t)$, one can write (1) as $x' = f(x; t)$. In particular, if $F = F(x; t)$ is any scalar function of class $C^{(1)}$ on the $(2n + 1)$ -dimensional $(x; t)$ -domain, then the total derivative of $F(x; t) = F(x(t); t)$ along any solution path $x = x(t)$ of (1) is $F' = F_t + F_x \cdot x' = F_t - F_x \cdot 1H_x \equiv F_t + (H; F)$, by (19), §20; so that $F' = \nabla F$, by (24₁), §21. This means that in case of Hamiltonian

systems one can replace Δ in (2), §79 by ∇ ; i.e., that, for any $F = F(x; t)$,

$$(2) \quad \Delta F = F' = F_t + (H; F) = \nabla F \text{ along solution paths } x = x(t) \text{ of (1).}$$

§92. According to (25), §21, the function $\nabla(F^1; F^2)$ of $(x; t)$ vanishes identically whenever the same holds for ∇F^1 and ∇F^2 , where both $F(x; t)$ are supposed to be of class $C^{(2)}$. It follows, therefore, from (2) and from the definition of an integral (§82), that if F^1 and F^2 are integrals of (1), then either the function $(F^1; F^2)$ of $(x; t)$ is a constant (which is, e.g., the case if F^1, F^2 are in involution; cf. §23), or else $\nabla(F^1; F^2)$ is again an integral of (1). In the latter case, $(F^1; F^2)$ may, but need not, be a new integral of (1), i.e., one which is independent of F^1, F^2 ; cf. §23–§24.

It is seen from (2) and §82 that a non-constant conservative function $F(x)$ of class $C^{(1)}$ is an integral of (1) if and only if it is in involution with the Hamiltonian function $H(x; t)$ for every fixed t . Since $(G; G) \equiv 0$ (cf. §20), it is also seen from (2) that if $H(x; t) \not\equiv 0$ (i.e., $f \not\equiv 0$; cf. §82), then $H(x; t)$ itself is an integral of (1) if and only if $H_t \equiv 0$. Thus, those Hamiltonian systems (1) which are conservative are characterized by the existence of the “energy integral”

$$(3) \quad H(x) = h, \quad \text{where } h = \text{const.} = H(x^0);$$

so that the integration constant h of the energy is a function of class $C^{(2)}$ of the $2n$ initial integration constants represented by $x^0 = x(t^0)$.

§93. A non-conservative Hamiltonian system (1) with n degrees of freedom can be replaced by a conservative Hamiltonian system

$$(4) \quad p'_j = -H_{q_j}(p, q), \quad q'_j = H_{p_j}(p, q), \quad (j = 0, 1, \dots, n),$$

with $n + 1$ degrees of freedom, where $q_j = q_i, p_j = p_i$ for $j = i > 0$ (cf. also §9 bis, §90). In order to see this, introduce the time as an $(n + 1)$ -th coordinate, and define a conservative $H(p, q)$ by placing

$$(5) \quad H(p, q) = H(p, q; q_0) + p_0; \quad \text{so that } q_0 = t,$$

while p_0 is, for the moment, arbitrary. Clearly, those equations (4) in which $j > 0$ are identical with (1); while those with $j = 0$ become $p'_0 = -H_t(p, q; t), q'_0 = 1$, i.e., $(H(p, q))' = 0, q_0 = t - \bar{t}$. Hence, the integration constant \bar{t} must be chosen as $\bar{t} = 0$; while $(H(p, q))' = 0$ is satisfied along any solution of the conservative system (4), since (4) has an energy integral $H(p, q) = h = \text{const.}$ Since one can add arbitrary constants to H, H in (1), (4), one may

choose $h = 0$; so that $H(p; q) = 0$. Then it is seen from (5) that the momentum p_0 canonically conjugate to the coordinate $q_0 = t$ is $p_0 = -H(p, q; t)$.

§94. Suppose that the Hamiltonian function of (1) has a non-vanishing n -rowed Hessian $\det(H_{p_i p_k}(p, q; t))$ in the $(2n + 1)$ -dimensional $(p, q; t)$ -domain. Then the Hamiltonian data $p, q, H(p, q; t)$ and $\det(H_{p_i p_k}) \neq 0$ become, in virtue of the point transformation of §15, equivalent to the Lagrangian data $q', q, L(q', q; t)$ and to $\det(L_{q'_i q'_k}) \neq 0$, respectively. Since (17), §19 is an identity in virtue of this point transformation, the Hamiltonian system $x' + 1H_x = 0$ for paths $x = x(t)$ in the $2n$ -dimensional phase space $x = (p, q)$ is equivalent to the Lagrangian system

$$(6) \quad [L]_q = 0,$$

$$\left([L]_{q_i} \equiv \sum_k q_k'' L_{q'_i q'_k} + \sum_k q_k' L_{q'_i q_k} + L_{q'_i t} - L_{q_i}; \text{ cf. §9} \right),$$

for paths $q = q(t)$ in the n -dimensional configuration space q . Correspondingly, the equivalent equations (1) and (6) are of first and second order, respectively.

Since $\det(L_{q'_i q'_k}(q', q; t)) \neq 0$ in the $(2n + 1)$ -dimensional $(q', q; t)$ -domain, one can solve (6) with respect to q'' ; so that, if z denotes the $2n$ -vector whose components are those of the n -vectors $r = q'$ and q together, one can write (6) in the form $z' = g(z; t)$; an equation to which §90 and what precedes §90 are applicable. Notice, however, that if (1), (6) are written as $x' = f(x; t)$, $z' = g(z; t)$, and if f is of class $C^{(\nu)}$ for some fixed $\nu \geq 1$, then g need not be of the same class $C^{(\nu)}$. This is particularly disagreeable in the limiting case $\nu = 1$, and shows that (1) must often be preferable to (6).

If $\det(H_{p_i p_k})$ or $\det(L_{q'_i q'_k})$ vanishes, then* the passage from (1) to (6) or from (6) to (1) is not defined by §15. The local existence theory (§79–§90) is applicable to (1) also when $\det(H_{p_i p_k}) = 0$, but not to (6) when $\det(L_{q'_i q'_k}) = 0$. Thus, the non-vanishing of one, hence of both, of these Hessians will be assumed whenever not merely (1) but also (6) is considered.

It should be mentioned that if $G(q)$ is any scalar function of class

* It is, however, known from the calculus of variations that there is no actual difficulty in the particular case $\lambda L(q', q) \equiv L(\lambda q', q)$, $\lambda > 0$, of $\det(L_{q'_i q'_k}) \equiv 0$, provided that the rank ($\leq n - 1$) of $\det(L_{q'_i q'_k})$ is $n - 1$ ("indicatrix" and "figuratrix").

$C^{(2)}$ in the configuration space, one can add to L in (6) not only any constant but also the linear form $G_q(q) \cdot q' \equiv (G(q))'$ of the q'_i , since $[G_q \cdot q']_q \equiv 0$ by the definition of $[]_q$.

§95. If $q = q(\bar{q}; t)$ is a coordinate transformation of the type considered in §10, and if $\bar{L}(\bar{q}', \bar{q}; t)$ is defined as there, (8), §10 shows that the Lagrangian equations are invariant. This holds, of course, only as long as $\det q_{\bar{q}} \neq 0$. An example of astronomical significance will in §343 (and, more generally, in §340–§342) show how wrong can be the results obtained, if one replaces $[L]_q = 0$ by $[\bar{L}]_{\bar{q}} = 0$ in case the n scalar equations which define a transformation $q = q(\bar{q}; t)$ or $q = q(\bar{q})$ are dependent, so that the Jacobian $\det q_{\bar{q}} \equiv 0$.

The invariance of the Lagrangian equations (6) under transformations $q = q(\bar{q}; t)$ of class $C^{[2]}$, and also the last remark of §94, become evident by observing that, as long as broken extremals are not considered, (6) is equivalent to the condition

$$(7) \quad \bar{\delta} \int L(q', q; t) dt = 0$$

for the extremals $q = q(t)$ of a calculus of variations problem with unvaried* boundaries.

§96. If, for a given $L(q', q; t)$, there is known a family of coordinate transformations which depend on a parameter ϵ , tend to the identical transformation as $\epsilon \rightarrow 0$, satisfy the differentiability conditions of §11 and are such as to leave $L(q', q; t)$ invariant in the sense of §11 bis or, at least, in the sense of §11, then the Lagrangian equations $[L]_q = 0$ possess the integral

$$(8) \quad f(q', q; t) \cdot L_{q'}(q', q; t) = \text{const.} \quad (\text{if } f \cdot L_{q'} \neq \text{Const.}),$$

where the n -vector function f is obtained by differentiating the transformation formulae with respect to ϵ at $\epsilon = 0$. In fact, (8) is clear from (11), §11, since $[L]_q = 0$.

§96 bis. Using (4), §9 instead of (11), §11, one sees that $[L]_q = 0$ has the integral

$$(9) \quad -L + q' \cdot L_{q'} = h = \text{const.} \quad (\text{if } -L + q' \cdot L_{q'} \neq \text{Const.}),$$

* This restriction is indicated by the dash of the variation symbol $\bar{\delta}$ in (7). In other words, $\bar{\delta}$ means that $t'(c)$ and $q(t'(c))$ in §14 are supposed to be independent of c .

if and only if $L_t \equiv 0$, i.e., $L = L(q', q)$. This contains, however, nothing new, since (9) is identical with the energy integral (3); cf. (1₁), (2₁), §15.

§97. Consider, as in §14, two functions $t^I(c)$, $t^{II}(c)$ and a family of paths $q = q(c; t)$ which satisfy the differentiability conditions of §14. Suppose further that the $m(\geq 1)$ components of the parameter vector $c = (c_j)$ are integration constants of $[L]_q = 0$, where $L = L(q', q; t)$; i.e., that $q = q(c; t)$ is, for every fixed c , a solution of the Lagrangian system $[L]_q = 0$. Then (19), §14 reduces, in view of (1₁)–(2₁), §15, to

$$(10) \quad \delta S = - (H)_{t=t^{II}} \delta t^{II} + (H)_{t=t^I} \delta t^I + \\ (p)_{t=t^{II}} \cdot \delta(q)_{t=t^{II}} - (p)_{t=t^I} \cdot \delta(q)_{t=t^I},$$

where, according to (18), §14 and (20), §14,

$$(11_1) \quad S \equiv S(c) = \int_{t^I(c)}^{t^{II}(c)} L(q'(c; t), q(c; t); t) dt; \quad (11_2) \quad \delta = \sum_{j=1}^m \frac{\partial}{\partial c_j} dc_j.$$

If, in particular, the system is conservative, then (9) holds along every solution path for an integration constant $h = H$ (which is, of course, a function $h = h(c)$ of the integration constants c_j); so that (10) reduces to

$$(12) \quad \delta S(c) = - h \delta t^{II} + h \delta t^I + \\ (p)_{t=t^{II}} \cdot \delta(q)_{t=t^{II}} - (p)_{t=t^I} \cdot \delta(q)_{t=t^I}; \quad h = h(c).$$

Notice that the integration constants c_j need not be independent; hence, their number, m , need not be less than a number depending on the degree of freedom, n .

In addition, use will be made of the fact that, by (11₂), one has $\delta f(c) = df(c)$ for a function f of the c_j alone, and so, in particular, for $f = c$.

§98. Suppose that the family $q = q(c; t)$ considered in §97 has the particular structure

$$(13) \quad q = q(c; t) \equiv q(q^0, \bar{q}, t^0, \bar{t}; t); \quad q^0 = (q)_{t=t^0}, \quad \bar{q} = (q)_{t=\bar{t}},$$

where $t^0 \leq t \leq \bar{t}$; so that the integration constants $q^0 = (q_i^0)$ and $\bar{q} = (\bar{q}_i)$ represent the “initial” and “final” positions in the configuration space along a solution path of the family, t^0 and \bar{t} being two

additional integration constants which will be considered as independent parameters.

According to (13), the m parameters c_j of §97 are represented by the $m = 2n + 2$ integration constants $q_i^0, \bar{q}_i, t^0, \bar{t}$. Hence, if one identifies t^0, \bar{t} with t^I, t^{II} , respectively, (11₁) becomes

$$(14) \quad S \equiv S(q^0, \bar{q}, t^0, \bar{t}) = \int_{t^0}^{\bar{t}} L(q', q; t) dt, \text{ where } q = q(q^0, \bar{q}, t^0, \bar{t}; t),$$

while (10) reduces, by the last remark of §97, to

$$dS = - (H)_{t=\bar{t}} d\bar{t} + (H)_{t=t^0} dt^0 + (p)_{t=\bar{t}} d\bar{q} - (p)_{t=t^0} dq^0.$$

This relation states that the partial derivatives of (14) are

$$(15_1) \quad S_{q^0} = - (p)_{t=t^0}, \quad S_{\bar{q}} = (p)_{t=\bar{t}};$$

$$(15_2) \quad S_{t^0} = (H)_{t=t^0}, \quad S_{\bar{t}} = - (H)_{t=\bar{t}}.$$

§99. If L is of the conservative type $L = L(q', q)$, then, by §79, only the difference $\bar{t} - t^0$ of \bar{t} and t^0 occurs in (13), while H has, by (3), a value h independent of t along any solution path; so that (13), (15₂) reduce to

$$(16_1) \quad q = q(q^0, \bar{q}, \bar{t} - t^0; t); \quad (16_2) \quad S_{t^0} = h, \quad S_{\bar{t}} = -h.$$

Substitution of (16₁) into (9) shows that the energy constant h is a function $h(q^0, \bar{q})$ of the integration constants. Actually, q^0 and \bar{q} are, by (13), two different positions in the configuration space along one and the same solution path; so that h is, with reference to (16₁), a function of q^0 alone:

$$(17) \quad h = h(q^0), \quad \text{where } q^0 = (q_i^0), \quad (i = 1, \dots, n).$$

If, on using (14) and (17), one defines a function W of the integration constants $q^0, \bar{q}, t^0, \bar{t}$ by placing

$$(18) \quad W = S + h(q^0)(\bar{t} - t^0) \equiv \int_{t^0}^{\bar{t}} (L + h) dt, \text{ then } W = W(q^0, \bar{q}),$$

i.e., W is independent of t^0 and \bar{t} . In fact, (16₂) shows that the partial derivative of the sum (18) with respect to t^0 or \bar{t} vanishes identically. Since S is thought of as expressed in terms of $q^0, \bar{q}, t^0, \bar{t}$, its partial derivative S_h with respect to (17) vanishes identically; and so (18) implies for the time elapsed between the positions q^0 and \bar{q} the representation

$$(19) \quad \bar{t} - t^0 = W_h(q^0, \bar{q}).$$

Furthermore, S is, by (18), a (linear) function of $\bar{t} - t^0$, and not of \bar{t} and t^0 separately. Finally, the integrand in (18) is, in view of the energy integral $H = h$ and of the definition $L = -H + p \cdot q'$ (§15), identical with $p \cdot q'$. Thus,

$$(20_1) \quad \int_{t^0}^{\bar{t}} L(q', q) dt = S \equiv S(q^0, \bar{q}, \bar{t} - t^0);$$

$$(20_2) \quad W(q^0, \bar{q}) = \int_{t^0}^{\bar{t}} p \cdot q' dt.$$

The content of (20₂) is that of expressing the line integral $\int p \cdot dq$ as a function of the end-points q^0, \bar{q} of the solution arc in the configuration space.*

§100. As another application† of §97, suppose that the given family of particular solutions $q = q(c; t)$ of a conservative Lagrangian system $[L]_q = 0$ consists of paths which are closed in the n -dimensional q -space, i.e., that $q(c; t + \tau) \equiv q(c; t)$ holds for every c and some period $\tau = \tau(c) > 0$. Suppose further that this function $\tau = \tau(c)$ of c is‡ of class $C^{(1)}$. Then τ is a single-valued function of the energy constant h alone; i.e., the period $\tau(c)$ does not depend on the m individual integration constants c_j which constitute $c = (c_j)$, but merely on their combination $h = h(c)$.

In order to prove this, notice first that, since $q(c; t)$ has the period $\tau \equiv \tau(c)$, the same holds for $q'(c; t)$, and so for $p = p(c; t)$ also (cf. (1₁), §15, where, by assumption, L does not contain t explicitly).

* The above remarks, together with §13 bis, form the formal basis of the theory of fields in calculus of variations. However, the content of the relations of §98–§99 is essentially less than that of the corresponding relations in the theory of fields. In fact, the relations of §98–§99 do not depend on the notion and on the existence of a field of extremals (Beltrami, Weierstrass, Poincaré, Hilbert) and are essentially older (Hamilton, Jacobi).

† The result of this article, often rediscovered in the mathematical literature, goes back to early efforts in classical statistical mechanics which attempted to find analogies between theorems of ordinary (i.e., non-statistical) mechanics and the second law of thermodynamics.

‡ This assumption is essential. It is not satisfied at a fixed $c = \bar{c}$ if $\tau(c)$ behaves at this \bar{c} the same way as $|c - \bar{c}|^{\frac{1}{2}}$ or as the third root of $(c - \bar{c})^2$, say. This explains why, in families of periodic solutions of the restricted problem of three bodies, the period is a single-valued function of the Jacobi constant (= energy) only locally and not in the large.

Hence, on choosing $t^{\text{II}}(c) = \tau(c)$ and $t^{\text{I}}(c) \equiv 0$ in (11₁), one sees from (11₂) that (12) reduces to $\delta S(c) = -h(c)\delta\tau(c)$. In fact, the second term on the right of (12) vanishes identically, while the third cancels the fourth. Since (11₂), when applied to functions of c alone, can be replaced by the symbol d of total differentiation in the c -domain, it follows that $dS = -hd\tau$. Since $d(h\tau) \equiv hd\tau + \tau dh$, one can write this as $dW = \tau dh$, if $W = W(c)$ denotes* the function $S + \tau h$ of c . Now, $dW(c) = \tau(c)dh(c)$ implies that if the m integration constants c_j which occur in $h = h(c)$ vary in such a way as to leave the value of $h(c)$ unchanged, then the value of $W(c)$ also remains unchanged. This means that W is a function of h alone. Hence, the same holds for the derivative of this function of h with respect to h . Since $dW = \tau dh$ shows that this derivative exists and is equal to the period τ , the proof is complete.

It is also seen that $\tau = \tau(h)$ is independent of h (i.e., that all solutions of the family have a common period) if and only if $W = W(h)$ is linear in h , in which case the same holds for $S(h) \equiv W(h) - \tau(h)h$.

§101. Assuming either, hence both, of the systems (1) and (6) to be conservative, one can apply §85 to any fixed solution $x = \bar{x}(t)$ of (1) or to the corresponding solution $q = \bar{q}(t)$ of (6). It is easily seen from the definitions (8)–(9), §85, that the Jacobi equations which determine the displacements of the solution $x = \bar{x}(t)$ or $q = \bar{q}(t)$ of (1) or (6) are again Hamiltonian or Lagrangian systems, respectively; namely,

$$(21_1) \quad I\xi' = \mathbf{H}_\xi, \quad \mathbf{H}(\xi; t) = \frac{1}{2}\xi \cdot H_{xx}(\bar{x}(t))\xi;$$

$$(21_2) \quad [\mathbf{L}]_\kappa = 0, \quad \mathbf{L}(\kappa', \kappa; t) = \frac{1}{2}\zeta \cdot L_{zz}(\bar{z}(t))\zeta.$$

It is understood that the $2n$ -matrices of the quadratic forms \mathbf{H} ; \mathbf{L} represent the Hessian matrices of the Hamiltonian and Lagrangian functions $H(x)$; $L(z)$ along the given solution, while $x = (x_j)$; $z = (z_j)$ denote the $2n$ -vectors defined by $x_i = p_i$, $x_{i+n} = q_i$; $z_i = q_i'$, $z_{i+n} = q_i$, finally ξ ; ζ displacements of $x = \bar{x}(t)$; $z = \bar{z}(t)$, respectively, while $\zeta = (\kappa', \kappa)$.

Furthermore, it is easily verified that the Hamiltonian and Lagrangian functions \mathbf{H} ; \mathbf{L} belong to each other in the sense of §15.

Since (1), where $H_t \equiv 0$, has the integral (3), it is clear from the

* While W is thus introduced *ad hoc*, §116 bis and §118 will show that W has a deeper and more general significance. Cf. also (18)–(20), §99.

rule (14), §87 that the Jacobi system belonging to a solution $x = \bar{x}(t)$ of (1) has the integral

$$(22) \quad \xi \cdot H_x(\bar{x}(t)) = h, \quad \text{where } h = \text{const.} \quad (\text{if } H_x(x) \neq 0).$$

§102. If a displacement of $x = \bar{x}(t)$, i.e., a solution $\xi = \xi(t)$ of (21₁), is such that its integration constant h defined by (22) vanishes, then $\xi = \xi(t)$ is called an isoenergetic displacement of $x = \bar{x}(t)$. What one actually means is that those displacements (that is, by §86, those infinitesimal displacements) for which the energy constant $h = H(\bar{x}(t))$ of the given solution satisfies $H(\bar{x}(t) + h\xi(t)) = h + o(|h|)$ are characterized by the vanishing of the constant h defined by (22); while* $O(|h|)$ instead of $o(|h|)$ holds for any displacement $\xi(t)$, i.e., for any solution $\xi(t)$ of (21₁) and for its integration constant (22). The proof of these statements is the same as the proof given at the end of §86. It is clear from §86 that, whether $h = 0$ or $h \neq 0$, the function $\bar{x}(t) + h\xi(t)$ is not, in general, a solution $x(t)$ of (1) (so that the error terms $o(|h|)$, $O(|h|)$ must be considered as functions of t also); but that the estimates $o(|h|)$, $O(|h|)$ hold uniformly on the t -interval $0 \leq t \leq M$ of §86.

The situation becomes clear by considering, as in §87, a family of solutions $x = x(t; \epsilon)$ of (1) which reduce at $\epsilon = 0$ to $x = \bar{x}(t)$. Clearly, the energy constant (3) becomes, within this family, a function $h = h(\epsilon)$ of the integration constant ϵ of (1). Now suppose that the family consists of isoenergetic solutions of (1), i.e., that $h = h(\epsilon)$ is independent of ϵ . Then the particular displacement $\xi(t)$ which is derived from $x(t; \epsilon)$ by the rule (13), §87 is an isoenergetic displacement of $\bar{x}(t) \equiv x(t; 0)$, since its integration constant h obviously vanishes. Actually, the same holds also when the derivative of $h(\epsilon)$ with respect to ϵ vanishes only at $\epsilon = 0$, and not for every ϵ .

Since the solutions $\bar{x}(t)$ and $\bar{x}(t + \epsilon)$ of (1) have, by (3), an energy constant h which is independent of ϵ , the displacement $\xi(t) = \bar{x}'(t)$ mentioned at the end of §87 is isoenergetic.

Solutions and Canonical Transformations

§103. The significance of the theory developed in §27–§46 lies in the fact that a phase space transformation of the type considered in §17 does or does not send *every* Hamiltonian system into a system

* By $O(\epsilon)$ and $o(\epsilon)$ are meant terms for which one respectively has $|O(\epsilon) : \epsilon| < \text{const.}$ and $|o(\epsilon) : \epsilon| \rightarrow 0$, as $\epsilon \rightarrow 0$.

which is again Hamiltonian according as the transformation is or is not canonical. This is clear from §26 bis, where (2), with $\Gamma = y_x$, is an identity for any transformation $y = y(x; t)$.

If $y = y(x; t)$ is canonical, then, by §27, the systems

$$(1_1) \quad Ix' = H_x(x; t);$$

$$(1_2) \quad Iy' = K_v(y; t), \quad \text{where} \quad K = \mu H(x(y; t); t) + R(y; t),$$

are equivalent for any $H(x; t)$, where the constant multiplier μ and the function R depend only on the transformation and not on the choice of $H(x; t)$. In fact, by §27,

$$(2_1) \quad \Gamma I \Gamma' = \mu I; \quad (2_2) \quad Iy_t = R_v, \quad \text{where} \quad y_t = y_t(y; t), \quad R = R(y; t),$$

while $\Gamma = \Gamma(y; t)$ is the Jacobian $2n$ -matrix y_x , and $y_t = y_t(y; t)$ the partial derivative defined in §17.

In what follows, use will be made of the formal observation that the necessary condition (2₂) for canonical transformations $y = y(x; t)$ would represent a Hamiltonian system, with R as Hamiltonian function, if one should replace the partial derivative $y_t = y_t(y; t)$ in the $(2n + 1)$ -dimensional $(y; t)$ -space by the total derivative $y' = y'(t)$ along a path in the phase space y . This fact is usually described by saying that the canonical transformations are contact transformations.

§104. Let $x = x(c; t)$ be a general solution of (1₁), in the sense that, in contrast with §83, the $2n$ integration constants c_i which constitute $c = (c_i)$ need not be initial values $x_i^0 = x_i(t_0)$ but are allowed to be arbitrary independent combinations of the latter. In other words, x^0 is replaced by an arbitrary $c = c(x^0)$ of non-vanishing Jacobian $\det c_{x^0}$. It is understood that $c = c(x^0)$, hence also the corresponding general solution $x = x(c; t)$, is supposed to satisfy the necessary differentiability conditions.

If the set c of $2n$ integration constants c_i of (1₁) happens to be such that the transformation of c into x , as defined by the corresponding general solution $x = x(c; t)$, is a canonical transformation of multiplier $\mu = 1$, then the c_i are called canonical integration constants of (1₁).

It turns out that this is the case if and only if the conservative transformation $x = x(c; t_0)$ which belongs to a suitable fixed t_0 is a canonical transformation of multiplier $\mu = 1$.

The necessity of this condition for a canonical set c of integration

constants is obvious from the first remarks of §36, which also show that if there exists one t_0 , then t_0 can be chosen arbitrarily. In order to prove the sufficiency of the condition, change the notations by using the letters y, x, R instead of x, c, H , respectively; so that (1₁) and its general solution $x = x(c; t)$ become

$$(3_1) \quad Iy' = R_v(y, t);$$

$$(3_2) \quad y = y(x; t).$$

Thus, $Iy'(x; t) = R_v(y(x; t), t)$ and $y'(x; t) \equiv y_t(x; t)$. Hence, if $x = x(y; t)$ is the inverse of (3₂) and one puts $R(y; t) = R(y(x(y; t); t), t)$ and $y_t(y; t) = y_t(x(y; t); t)$, it is clear from the differentiation agreements of §17 that condition (2₂) of §103 is satisfied. In other words, condition (ii) of §36 is satisfied. On the other hand, condition (i) of §36 requires the existence of a t_0 for which the conservative transformation $y = y(x; t_0)$ is canonical. Since this condition is satisfied by assumption, the proof is complete.

§104 bis. Since the transition from the initial values of a Hamiltonian system to any of its canonical sets of integration constants is, by §104, a conservative canonical transformation of multiplier $\mu = 1$, it is, by §35, a completely canonical transformation (i.e., one which does not contribute anything to the new Hamiltonian function; cf. §34).

§105. If $x = x(x^0; t)$ is the general solution of any fixed canonical system (1₁) in terms of the $2n$ initial values x_i^0 which are assigned to a $t = t_0$, then $y = y(x; t)$, where $y = x^0$, is a canonical transformation which has the multiplier $\mu = 1$ and a remainder function R which is identical with the negative of the Hamiltonian function H of (1₁).

In fact, on applying the criterion of §104 to $c = x^0$, and noting that $x = x(x^0; t_0)$ is the identical transformation $x = x^0$ (which is a canonical transformation of multiplier $\mu = 1$), one sees that the x_i^0 are canonical integration constants of (1₁). This means that $y = y(x; t)$, where $y = x^0$, is a canonical transformation, with $\mu = 1$. Since the remainder function $R(y; t)$ of $y = y(x; t)$ depends only on the transformation $y = y(x; t)$ and not on the canonical system to which it is applied, one can determine R by applying $y = y(x; t)$ to a particular Hamiltonian system. Choosing the latter system as the system (1₁) whose general solution is the inverse of $y = y(x; t)$, where $y = x^0$, one sees from $x^0 = \text{const.}$ and $\mu = 1$ that (1₂) reduces to $0 = H_v + R_v$. Since this is, by §27, equivalent to $R = -H$, the proof is complete.

§105 bis. Since $\mu = 1$ implies, by §32, that $\det \Gamma = 1$, the mapping $x = x(x^0; t)$ of $x^0 = x(t_0)$ on $x = x(t)$ in the $2n$ -dimensional phase space is, for every fixed t , not only volume preserving but orientation preserving as well, no matter what is the Hamiltonian function $H(x; t)$ of (1₁). On the other hand, $x = x(x^0; t)$ is, by §34, completely canonical only when $R \equiv 0$. And $R \equiv 0$ means, in view of $R = -H$, that $H \equiv 0$, i.e., that (1₁) degenerates into the trivial system for which every solution is an equilibrium solution (cf. §82).

§106. Suppose that $I\mathbf{x}' = \mathbf{H}_{\mathbf{x}}(\mathbf{x}; t)$ is a given Hamiltonian system for which one knows the general solution in terms of a set $x = (x_i)$ of canonical integration constants x_i . Then the Hamiltonian function $K = K(y; t)$ of the system $Iy' = K_y(y; t)$ into which an arbitrary Hamiltonian system $I\mathbf{x}' = H_{\mathbf{x}}(\mathbf{x}; t)$ of the same degree of freedom is transformed by $y = \mathbf{x}(x; t)$ is $K = H + \mathbf{H}$.

In fact, $\mu = 1$; so that the statement $K = H + \mathbf{H}$ is, in view of (1₂), equivalent to the statement that the remainder function of $\mathbf{x} = \mathbf{x}(x; t)$ is \mathbf{H} , or (what is, by §31, the same thing) that the remainder function of $x = x(\mathbf{x}; t)$ is $-\mathbf{H}$. But this is clear from §105 (and §104 bis), since $\mathbf{x} = \mathbf{x}(x; t)$ denotes the general solution of $I\mathbf{x}' = \mathbf{H}_{\mathbf{x}}(\mathbf{x}; t)$ in terms of its set $x = (x_i)$ of canonical integration constants.

§107. Clearly, one can read the rule of §106 also in a reverse direction, as follows:

If one knows the general solution $\mathbf{x} = \mathbf{x}(x; t)$ of a particular Hamiltonian system $I\mathbf{x}' = \mathbf{H}_{\mathbf{x}}(\mathbf{x}; t)$ in terms of $2n$ canonical integration constants x_i , then any Hamiltonian system $Iy' = K_y(y; t)$ is sent by the transformation $y = \mathbf{x}(x; t)$ into the Hamiltonian system $I\mathbf{x}' = H_{\mathbf{x}}(\mathbf{x}; t)$ whose Hamiltonian function is given by

$$(4) \quad H(x; t) = K(\mathbf{x}(x; t); t) - \mathbf{H}(\mathbf{x}(x; t); t).$$

This result is the celebrated rule for the "variation of (canonical) integration constants" in the theory of perturbations.

§108. The theory of partial differential equations of the first order (Cauchy, Hamilton and Jacobi) associates every given Hamiltonian system (1₁), where $H(x; t) = H(p, q; t)$ and $x_i = p_i$, $x_{i+n} = q_i$; ($i = 1, \dots, n$), with the differential equation

$$(5) \quad S_t + H(S_q, q; t) = 0; \quad (q = (q_i); i = 1, \dots, n),$$

where the scalar function $S = S(q, t)$ has to be determined as a solu-

tion of (5), i.e., in such a way that (5) becomes an identity in the $(n + 1)$ -dimensional $(q; t)$ -domain under consideration. The solution paths of (1_1) are the "characteristics" of the associated partial differential equation (5), which contains only the partial derivatives $S_t, S_{q_1}, \dots, S_{q_n}$ of the unknown function S , and not S itself.*

If v_1, \dots, v_n are n integration constants and

$$(6) \quad S = S(t, q; v), \quad \text{where } q = (q_i), \quad v = (v_i); \quad (i = 1, \dots, n),$$

is, for fixed v , a solution of (5), then (6) is called a complete solution of (5) in case (6) is of class $C^{(2)}$ in the $(2n + 1)$ -dimensional $(t, q; v)$ -domain and, in this domain, the n -rowed determinant

$$(7) \quad \det (S_{q_i v_k}(t, q; v)) \neq 0; \quad (S_{q_i v_k} = S_{v_k q_i}; \quad i, k = 1, \dots, n).$$

§109. If, starting with any given complete solution (6) of (5), one puts

$$(8) \quad \begin{aligned} S_q(t, q; v) &= p \\ S_v(t, q; v) &= -u \end{aligned} \quad \text{and} \quad \begin{pmatrix} p \\ q \end{pmatrix} = x, \quad \begin{pmatrix} u \\ v \end{pmatrix} = y,$$

then the components $y_i = u_i, y_{i+n} = v_i$ of the $2n$ -vector y constitute a set of canonical integration constants of (1_1) .

In fact, whether (6) does or does not satisfy (5) for a fixed v , the relations (7), (8) are identical with the assumptions (22), (23) of §46. Hence, (8) defines a canonical transformation $y = y(x; t)$ which has the multiplier $\mu = 1$ and the remainder function $R = S_t$; cf. (21), §46. Now, since (6) is supposed to satisfy (5) for fixed v , it is seen from the first of the relations (8) that $S_t + H = 0$, where $H = H(p, q; t)$. Since $R = S_t, \mu = 1$, it follows that $\mu H + R = 0$, i.e., that the Hamiltonian function $K = K(y; t)$ of the system (1_2) into which the system (1_1) is transformed by $y = y(x; t)$ vanishes identically. In other words, the transformation $y = y(x; t)$ is such that $Iy' \equiv 0$, i.e. $y'(t) \equiv 0$, holds along every solution path $x = x(t)$ of (1_1) . Thus, the $2n$ components of y are integration constants of (1_1) , and so canonical integration constants, the transformation $y = y(x; t)$ being canonical and of multiplier $\mu = 1$.

§110. The result, thus proved, is two-fold, In fact, it states that

* One must not confuse the (almost always non-linear) partial differential equation (5) of the first order in $n + 1$ independent variables t, q_i with the linear partial differential equation $\nabla F \equiv F_t + (H; F) = 0$ in $2n + 1$ variables $t, x_i = p_i, x_{i+n} = q_i$ which determines the integrals $F(x; t)$ of (1_1) ; cf. (2), §91 with §82.

if (6) is a complete solution of (5) and $y = y(x; t)$ denotes the locally topological transformation which is, in view of (7), implicitly defined by (8), then

(i) the components $y_i(x; t)$ of $y = y(x; t)$ represent $2n$ independent integrals of (1₁) or, what is the same thing, $x = x(y; t)$ is the general solution of (1₁) in terms of $2n$ independent integration constants y_i ; and

(ii) these y_i form a *canonical* set of integration constants for (1₁).

Clearly, (i) cannot be considered a result of practical value for the problem of integration, since it is hardly easier to find a complete solution of the partial differential equation (5) than to find the general solution of the system (1₁) of ordinary differential equations.* Accordingly, the actual merit of the result lies not in (i) but rather in (ii), i.e., in a rule which, in case a complete [or general] solution of (5) [or (1₁)] is known, supplies a procedure for finding different combinations of the integration constants which form canonical sets of integration constants; sets to which the fundamental rule of §107 is applicable.

§111. Consider the family of solution paths described at the beginning of §98; so that any particular solution path of (1₁) is characterized, in this family, by the initial and final states in the configuration space. Writing q, t instead of \bar{q}, \bar{t} , and choosing $t^0 = 0$, one can write the definition (14), §98 as

$$(9) \quad S(q, t; q^0) = \int_0^t L \, d\bar{t},$$

where it is understood that the integration is extended along a solution path between the initial and final states, $(q^0; 0)$ and $(q; t)$, in the configuration space. One calls (9) the action integral (with reference to the given family).

Now, (9) is a solution $S(q, t)$ of (5), with the components q_i^0 of q^0 as integration constants. In fact, the identities (15₁)–(15₂), §98 reduce, in the present notations, to

* Actually, the only proof known to-day for the existence of a complete solution of (5) is based on the existence of the general solution of (1₁). Furthermore, this existence proof for (5), supplied by Cauchy's theory of characteristics, presupposes that $H(x; t)$ is of class $C^{(2)}$, while nothing seems to be known about (5) when $H(x; t)$ is only of class $C^{(1)}$; not even when $H_x(x; t)$ satisfies an additional condition of the Lipschitz type. On the other hand, such assumptions on H are known to be sufficient for a treatment of (1₁).

$$(10_1) S_q = p; \quad (10_2) S_t = -H(p, q; t); \quad (10_3) S_{q^0} = -p^0.$$

And substitution of (10₁) into (10₂) shows that (9) satisfies (5) for every fixed q^0 .

§112. If, in particular, one knows that the family of solution paths which underlies the action (9) is so chosen that (7) is satisfied by $v = q^0$, it follows that the partial differential equation (5) possesses a complete solution, a solution postulated in §109. Now, the (local) existence of solution paths for which (9) satisfies the completeness condition

$$(11) \quad \det (S_{q_i q_k^0}(q, t; q^0)) \neq 0 \quad (i, k = 1, \dots, n)$$

can be proved* on the assumption that $H(x; t)$ is of class $C^{(2)}$. The existence of families of solution paths for which (9) satisfies (11) is identical with the (local) existence of fields in calculus of variations; so that the standard construction involved will be omitted. This the more as the existence of complete solutions of (5) will be used only for the purpose explained at the end of §110, hence only in cases in which a complete solution is available explicitly (cf., e.g., §214 and §221, §248).

§113. Since comparison of (10₁), (10₃) with (8) gives $p^0 = u$, $q^0 = v$, the result of §105 can be interpreted as that particular case of the result of §109 for which the complete solution (6) of (5) is an action (9) which satisfies (11). Actually, the idea in §109 is identical with that in §105, since it consists both times of the choice of a canonical transformation which sends a given Hamiltonian system (1₁) into a Hamiltonian system (1₂) whose Hamiltonian function $K(y; t) = K(u, v; t)$ vanishes identically.

Sometimes (cf., e.g., §117) it is convenient to replace this normal form $K(u, v; t) \equiv 0$ of an H by the less drastic normal form in which K is allowed to be an arbitrary function of the coordinates v_i and of t but does not contain the momenta u_i , where $i = 1, \dots, n$. The integration problem of (1₂) is of a trivial type in this case also. In fact, (1₂) can then be written as

$$(12) \quad v' = K_u(v; t) \equiv 0, \quad u' = -K_v(v; t);$$

so that, if $v^0 = (v_i^0)$, $u^0 = (u_i^0)$ represent $n + n$ arbitrary integration constants, the general solution $u = u(t)$, $v = v(t)$ of (1₂) becomes

* Cf. the footnote to §110.

$$(13) \quad v = v^0, \quad u = u^0 - \int_{t^0}^t K_v(v^0; \bar{t}) d\bar{t} \equiv u^0 - \left(\int_{t^0}^t K(v^0; \bar{t}) d\bar{t} \right)_{v^0},$$

and requires, therefore, only quadratures.

Since $K(u, v; t) \equiv 0$ is sufficient in order that (1₂) has the form (12), one can transform (locally) every Hamiltonian system (1₁) into a system of the trivial form (12). But the determination of a canonical transformation of this type does not differ from the problem of integration of (1₁); cf. §110.

§114. It will be assumed in what follows that the given Hamiltonian function $H(x; t)$ is of the conservative type; so that (1₁), (5) reduce to

$$(14_1) \quad p' = -H_q(p, q), \quad q' = H_p(p, q); \quad (14_2) \quad S_t + H(S_q, q) = 0.$$

If h is any fixed constant, and $W = W(q; h)$ any solution of the partial differential equation

$$(15) \quad H(W_q, q) = h$$

in the n -dimensional q -domain, then

$$(16) \quad S = -ht + W$$

clearly is a solution of (14₂). It is also clear that if $S = S(q, t)$ is a given solution of (14₂), the function W defined by (16) is a solution of (15); furthermore, it can be shown that the function W , thus defined, is independent of t , i.e., that every solution $S(q, t)$ of (14₂) is linear in t (cf. §115; also §99, §111).

If one subjects (14₁) to the canonical transformation which represents the canonical extension of a given transformation $\bar{q} = \bar{q}(q)$ in the configuration space, then the Hamiltonian function and the momenta respectively transform as an invariant and as the components of a covariant vector in the configuration space; cf. §48. Since the gradient W_q of a function $W = W(q)$ also transforms as a covariant vector, it follows that the correspondence between (15) and (14₁) is preserved by any coordinate transformation $\bar{q} = \bar{q}(q)$ and its canonical extension.

§115. If $S = S(t, q; v)$ is a complete solution of (14₂), then it is a linear function of t , i.e., the function W defined by (16) is independent of t .

In fact, if $S(t, q; v)$ is a complete solution of (14₂), then (8) defines, by §109, the general solution of (14₁) in terms of integration constants u, v . Since substitution of S_q from (8) into (14₂) show that $-S_t$ is identical with $H(p, q)$, and since (14₁) has the energy integral $H(p, q) = \text{const.}$, it is clear that S_t cannot contain t explicitly. This, when combined with (16), completes the proof of $W_t \equiv 0$.

It is also seen that the fixed constant h occurring in (15) has to be identified with the energy constant $H(p, q) = \text{const.}$ of (14₁).

§116. If v_1, \dots, v_n are n integration constants and if

$$(17) \quad W = W(q, v), \text{ where } q = (q_i), v = (v_i), \quad (i = 1, \dots, n),$$

is a function of class $C^{(2)}$ in the (q, v) -domain under consideration, then (17) is called a complete solution of (15) if, on the one hand,

$$(18) \quad \det (W_{q_i v_k}(q, v)) \neq 0; (W_{q_i v_k} = W_{v_k q_i}; i, k = 1, \dots, n),$$

and, on the other hand, the expression $H(W_q, q)$ on the left of (15) is made by (17) a function of v alone. Accordingly, the constant h occurring in (15) is made by (17) a function

$$(19) \quad h = h(v)$$

of the n integration constants v_i occurring in (17).

It is clear from §115 that (16), together with (19), establishes a reciprocal correspondence between the complete solutions (6), (17) of (5), (15), respectively. In fact, the respective completeness conditions (7), (18) are equivalent, since

$$(20) \quad S_{q_i v_k} \equiv W_{q_i v_k},$$

by (16), where $h_q \equiv 0$.

It follows that if (17) is any complete solution of (15), then

$$(21) \quad p = W_q(q, v), \quad u = h_v(v)t - W_v(q, v)$$

implicitly defines, for fixed t , a locally topological correspondence between (p, q) and (u, v) in such a way that $p = p(t), q = q(t)$ becomes the general solution of (14₁) in terms of canonical integration constants u, v . This is clear from §109, if one observes that (21) is identical with (8) in virtue of (16) and (19).

§116 bis. Suppose, in particular, that one of the n integration constants v_i occurring in a given complete solution (17) of (15) happens

to be the energy constant h ; so that $h = v_n$, say. Thus, $h_{v_n} = 1$, while $h_{v_l} = 0$ for $l < n$; cf. (19). Substituting this into (21), denoting the integration constant u_n by t^0 , and writing P_i, Q_i for u_i, v_i , respectively, one sees that

$$(22) \quad \begin{aligned} p_i &= W_{q_i}, & i &= 1, \dots, n; \\ P_l &= -W_{q_l}, & l &= 1, \dots, n-1; & t - t^0 &= W_h \end{aligned}$$

where $W = W(q_1, \dots, q_n, Q_1, \dots, Q_n)$, is an implicit representation of the general solution of (14₁) in terms of $2n$ canonical integration constants

$$(23) \quad P_1, P_2, \dots, P_{n-1}, \quad P_n \equiv t^0; \quad Q_1, Q_2, \dots, Q_{n-1}, \quad Q_n \equiv h.$$

§117. According to §110, the point in the rule (21) or in its particular case (22) is that of supplying *canonical* sets of integration constants. In those applications for which this point is not essential, it is often useful to utilize the knowledge of a complete solution of (15) in a slightly different way, as follows:

If $W = W(q, \omega)$ is any scalar function of two n -vectors $q = (q_i)$, $\omega = (\omega_i)$ which is of class $C^{(2)}$ and such that $\det (W_{q_i \omega_k}) \neq 0$, then

$$(24) \quad p = W_q(q, \omega), \quad \chi = W_\omega(q, \omega)$$

defines a completely canonical transformation of (p, q) into (ω, χ) . This is clear from the general rule (20)–(21) of §46, if one puts $u = \omega$, $v = \chi$ and chooses $S = W$; so that $S_t \equiv 0$. Accordingly, (24) transforms every system (14₁) into

$$(25) \quad \omega' = -K_\chi, \quad \chi' = K_\omega,$$

where $K = K(\omega, \chi) \equiv H(p, q)$ in virtue of (24). Hence, $K(\omega, \chi) \equiv H(W_q(q, \omega), q)$ in virtue of $\chi = W_\omega(q, \omega)$. Consequently, if the given function $W(q, \omega)$ is a complete solution (17), with the components ω_i of ω as n integration constants v_i , then $K(\omega, \chi) \equiv h(\omega)$ in view of (15) and (19). Hence, (25) reduces to $\omega' = 0$, $\chi' = h_\omega(\omega)$ and has, therefore, the general solution

$$(26) \quad \omega = \omega^0, \quad \chi = \nu t + \chi^0, \quad \text{where } \nu = h_{\omega^0}(\omega^0) = \text{const.},$$

and the components of the two constant n -vectors ω^0, χ^0 are the $2n$ integration constants (these are not canonical integration constants, since the canonical conjugate of ω is $\chi = \nu t + \chi^0$).

This is the desired normal form* of the general solution of (14₁). It is understood that the last remark of §113 applies again. In fact, (25)–(26) is a particular case of (12)–(13), if one interchanges the rôle of the coordinates v, χ and of the momenta u, ω .

§118. It is clear from §114–§116 that the considerations of §111 remain valid if one replaces (5) by (15) and, correspondingly, §98 by §99. Then v in (17), (19) is particularized to $v = q^0$, while (9) corresponds to†

$$(27) \quad W(q, q^0) = ht + S(q, t; q^0) \equiv \int_0^t p \cdot q' d\bar{t}, \text{ where } h = h(q^0);$$

cf. (16) in §114 and (17), (18), (20₂) in §99.

If, in particular, the value of h in (15) is preassigned, (19) shows that the energy constant $h = h(v) \equiv h(q^0)$ of the solution paths which constitute a family considered in §111 must be chosen independent of q^0 . Then (27) is called the isoenergetic action belonging to this isoenergetic family of solutions.

Non-local Notions

§119. The preceding notions and considerations are of a local nature. This remark also applies to §84, where it was supposed, instead of being proved, that the particular solution $x = x(t)$ of

$$(1) \quad x' = f(x), \quad (x = (x_i), f = (f_i), i = 1, \dots, m),$$

exists on a t -interval of arbitrarily large but fixed length $M < +\infty$. The notions which will now be considered concern problems on the infinite t -axis, problems in the large for which there clearly cannot exist a general theory of the type of §79 or §84.

A particular solution $x = x(t)$ of (1) will be called unrestricted if it exists for $-\infty < t < +\infty$. It is understood that $x(t)$ must lie for every t in the x -domain \mathbf{X} on which the function $f(x)$ of class $C^{(\nu)}$, where $\nu \geq 1$, is given, and that $x(t)$ must not cease to have a

* This normal form, introduced by Poincaré, is useful, for instance, for the treatment of formal trigonometrical expansions in celestial mechanics.

The coordinates χ_i and the momenta ω_i are, up to a trivial normalization, the angular coordinates and canonically conjugate action momenta in classical quantum theory ($\chi = w, \omega = J$ in the standard notations of the physicists).

† In this connection, cf. the last of the $2n$ relations (22), §116 bis with (19), §99.

finite derivative $x'(t)$ at any $t = \bar{t}$ (since otherwise $x = x(t)$ would not satisfy (1) at $t = \bar{t}$). For instance, if $m = 1$, $f(x) = x^2$ and \mathbf{X} : $-\infty < x < +\infty$, then only the equilibrium solution $x(t) \equiv 0$ is an unrestricted solution of (1), since every other solution is of the form $x(t) = (\bar{t} - t)^{-1}$. If $f(x) = x^2$ is replaced by $f(x) = 1$, every solution of (1) is unrestricted, though not bounded.

If $x(t)$ is an unrestricted solution, then so is $x(t - t^0)$ for every $t^0 = \text{const.}$, and will not be considered as distinct from $x(t)$. If two unrestricted paths have at least one point of the domain \mathbf{X} in common, then the two paths are identical, in view of the uniqueness of the local initial value problem of (1). It is understood that by an unrestricted path is meant the set of points $x = \bar{x}$ which can be parametrized by means of an unrestricted solution $x(t)$ in the form $\bar{x} = x(t)$, $-\infty < t < +\infty$. An unrestricted path need not be a closed set in \mathbf{X} . It is clear that every equilibrium point (§83) is an unrestricted path; and that every unrestricted path is an invariant set (§81).

§120. Every set \mathbf{X}^* of points x which consists of a (finite or infinite) collection of unrestricted paths is an invariant set. Any set \mathbf{X}^* of this type will be called an unrestricted invariant set of (1), it being understood that \mathbf{X}^* must be a subset of the x -domain \mathbf{X} on which $f(x)$ is given.

If (1) is given, for $m = 2$, as $x_1' = 1/x_2$, $x_2' = 1$, where \mathbf{X} : $-\infty < x_1 < +\infty$, $0 < x_2 < +\infty$, the general solution of (1) is seen to be $x_1(t) = \log(t - \bar{t}) + \text{const.}$, $x_2(t) = t - \bar{t}$; so that no solution is unrestricted, and hence \mathbf{X} cannot contain an unrestricted invariant set. If, on the other hand, $m = 1$, $f(x) = 1$, \mathbf{X} : $-\infty < x < +\infty$, then the domain \mathbf{X} itself is an unrestricted invariant set \mathbf{X}^* . This implies that an \mathbf{X}^* need not be a bounded set. Also when an \mathbf{X}^* is bounded, it need not be compact, i.e., closed.

An obvious adaptation of the considerations of §84 shows that if a subset \mathbf{X}^+ of \mathbf{X} is compact (i.e., such that the Heine-Borel theorem is applicable on \mathbf{X}^+), and if \mathbf{X}^+ is an invariant set (§81), then \mathbf{X}^+ is an unrestricted invariant set \mathbf{X}^* .

§121. For any given unrestricted invariant set \mathbf{X}^* of (1) and for every real t , one can define a one-to-one transformation τ^t of \mathbf{X}^* into itself, by placing $x(t) = \tau^t x^0$, where x^0 is any point of \mathbf{X}^* and $x(t)$ is that solution of (1) for which $x(t^0) = x^0$. In fact, the point $\tau^t x^0$

of \mathbf{X}^* , thus defined, is independent of the choice of t^0 (cf. §119). It is also clear from the existence and uniqueness of the solutions of (1), that $\tau^{t_1}x(t_2) = \tau^{t_2+t_1}x^0$ for arbitrary t_1, t_2 and for any point x^0 of \mathbf{X}^* . This means that $\tau^{t_1}\tau^{t_2} = \tau^{t_1+t_2}$. Thus, the transformations τ^t of \mathbf{X}^* which belong to different values of t form a (cyclic) group. In fact, on placing $t_1 = t, t_2 = -t$ and noting that τ^0 clearly is the identical transformation of \mathbf{X}^* into itself, one sees that τ^t has τ^{-t} as inverse transformation; cf. §79.

If $\bar{\mathbf{X}}$ is a set of points \bar{x} of \mathbf{X}^* , let $\tau^t\bar{\mathbf{X}}$ denote the set of all points $\tau^t\bar{x}$ for a given t . Thus, $\bar{\mathbf{X}}$ is an invariant set of (1) in the sense of §81 if and only if $\tau^t\bar{\mathbf{X}} = \bar{\mathbf{X}}$ for every t . Notice that an invariant set $\bar{\mathbf{X}}$ which consists of a single point \bar{x} represents the equilibrium solution $x(t) \equiv \bar{x} = \text{const.}$; cf. §83.

§121 bis. In the above and following considerations, it is permissible to think of the x -domain of $x' = f(x)$ not as a set contained in the m -dimensional Euclidean x -space but rather as a space which has a Euclidean structure only locally and not in the large.

Suppose, for instance, that the given domain of the m -vector function $f(x)$ is the whole Euclidean space, and that there exist an $r \leq m$ and r positive constants π_1, \dots, π_r in such a way that $f(x) \equiv f(x_1, \dots, x_m)$ remains unchanged if one replaces x_j by $x_j + \pi_j$, where $j = 1, \dots, r$. Then x_1, \dots, x_r may be thought of not only as linear variables but also as angular variables which have to be reduced mod $\pi_1, \dots, \text{mod } \pi_r$, respectively. If, in particular, $r = m$, one has $m + 1$ distinct choices as to the topological structure of the domain \mathbf{X} in the large, the two extreme choices being a Euclidean space and a torus. This holds for arbitrary π_1, \dots, π_r if $f(x)$ is independent of x , i.e., if

$$(2) \quad x' = \lambda, \text{ where } \lambda = (\lambda_i) = \text{const.}; \text{ so that } x(t) = \lambda t + x_0.$$

Corresponding remarks hold concerning the reduction of \mathbf{X} by means of any discontinuous group of transformations under which the function on the right of (1), §119 happens to be invariant.

Another extension of §119–§121 is obtained by allowing, when otherwise admissible, \mathbf{X} to consist not of a domain in the sense of §2 but of an open set and possibly of some or all of the boundary points of this open set. In such cases, \mathbf{X} will be called a region (so that every domain is a region).

§122. Of particular interest are the systems (1) for which the

m -rowed Jacobian (4), §79, which will be denoted by $D(x^0; t)$, is independent of x^0 and t . These particular systems (Liouville) may be called of the volume preserving type. In fact, if $D(x^0; t) \equiv \text{const.}$, then $D(x^0; t) \equiv 1$, as seen by placing $t = 0$ in (4), §79. If the general solution (3), §79 of a system $x' = f(x)$ is thought of as defining a "flow" in the x -space, the condition $D(x^0; t) \equiv 1$ defines the "incompressible" flows.

Although (4), §79 defines $D(x^0; t)$ in terms of the general solution (3), §79 of $x' = f(x)$, one can decide without any knowledge of the solutions of $x' = f(x)$, whether or not $x' = f(x)$ is of the volume preserving type. For to this end one merely has to calculate the diagonal elements of the Jacobian of the given function $f(x)$, and then see whether or not $\text{div } f(x) \equiv 0$. In fact, it is clear from (4 bis), §79, that the determinant (4), §79 is independent of t for every x^0 if and only if $\text{div } f$ vanishes identically.*

If $x' = f(x)$ is a canonical system with $n = \frac{1}{2}m$ degrees of freedom, then the condition $D(x^0; t) \equiv 1$ is satisfied, by §105 bis. Correspondingly, $\text{div } f(x) \equiv 0$ then is satisfied, since the components of the $2n$ -vector $f(x)$ are of the form $-H_{x_{k+n}}(x)$, $H_{x_k}(x)$, where $k = 1, \dots, n$.

§122 bis. It may be mentioned that if $m = 2$ (and only in this case), the condition $\text{div } f(x) \equiv 0$ is not only necessary but also sufficient for a system $x' = f(x)$ to be canonical. For if u, v and g, h denote the components of the 2-vectors x and $f(x)$, respectively, then $\text{div } f = g_u + h_v$; so that $\text{div } f(x) \equiv 0$ is precisely the integrability condition for the existence of a scalar $H = H(x)$ such that $g = -H_v$, $h = H_u$.

This fact implies Jacobi's principle of the last multiplier in the volume preserving case. For suppose that there are known $n - 2$ independent conservative integrals $F_1(x), \dots, F_{n-2}(x)$ of $x' = f(x)$. Let $\mathbf{R} = \mathbf{R}(c_1, \dots, c_{n-2})$ denote the (two-dimensional) intersection of the corresponding hypersurfaces $F_1(x) = c_1, \dots, F_{n-2}(x) = c_{n-2}$, where the integration constants c have fixed values. Since the lat-

* As an application, one can infer from (25), §232 that if $x' = f(x)$ is given by (24), §232, where $m = 3$, then $D(x^0; t) \equiv 1$.

In this connection, there arises the question, when is a given three-dimensional incompressible flow an isoenergetic flow belonging to a (conservative) canonical system with two degrees of freedom. This problem, when properly specified, and its higher-dimensional analogue are to-day unsolved; they appear to lead to analytico-topological questions, which are always rather difficult.

ter may be chosen arbitrarily, it is easy to see that the projection on \mathbf{R} of the n -dimensional incompressible flow of the x -space is again incompressible.† Accordingly, the flow on \mathbf{R} is defined by two scalar differential equations of the form $u' = g(u, v)$, $v' = h(u, v)$, where $g_u + h_v \equiv 0$. But $g_u + h_v \equiv 0$ means that $u' = g$, $v' = h$ is a conservative canonical system with a single degree of freedom, and may therefore be solved, in view of its energy integral, by a single quadrature.

Notice that these considerations are purely local in nature.

§123. Let $\mu(\mathbf{S})$ denote the volume measure of a Lebesgue measurable subset \mathbf{S} of the m -dimensional Euclidean space x of (1). Suppose that a given unrestricted invariant set \mathbf{X}^* of (1) has a measure $\mu(\mathbf{X}^*)$ which is neither 0 nor $+\infty$. Suppose further that (1) satisfies the condition $\operatorname{div} f(x) \equiv 0$ of §122 for the preservation of the measure μ ; so that, in the notations of §121, one has $\mu(\tau^t \mathbf{S}) = \mu(\mathbf{S})$, $-\infty < t < +\infty$, for every measurable subset \mathbf{S} of \mathbf{X}^* . Then the set of those points x^0 of \mathbf{X}^* for which the path $x = x(t) \equiv \tau^t x^0$ in \mathbf{X}^* does not possess an asymptotic distribution function $\psi_{x^0} = \psi_{x^0}(\mathbf{S})$ is a set of vanishing volume, i.e., of μ -measure 0. This is (or, rather, is equivalent to) the celebrated Ergodic Theorem of G. D. Birkhoff, which, as a matter of fact, has nothing to do with differential equations, since it represents a theorem belonging to the general theory of Lebesgue measure. Thus, the proof would be out of place in this book.

The formulation of the theorem, given above for the Euclidean case under consideration, depends on the notion of asymptotic distribution functions, which are defined as follows:

By a distribution function $\phi = \phi(\mathbf{S})$ on \mathbf{X}^* is meant a set-function which assigns to every Borel set \mathbf{S} (e.g., to every open set \mathbf{S} and to every closed set \mathbf{S}) contained in \mathbf{X}^* a real non-negative value $\phi(\mathbf{S})$ in such a way that $\phi(\mathbf{S}_1) + \phi(\mathbf{S}_2) + \dots = \phi(\mathbf{S}_1 + \mathbf{S}_2 + \dots)$ whenever the sets $\mathbf{S}_1, \mathbf{S}_2, \dots$ are mutually disjoint, while $\phi(\mathbf{X}^*) = 1$. It is known that if the discontinuity sets \mathbf{D} of a distribution function ϕ on \mathbf{X}^* are defined by the condition $\phi(\mathbf{D}^0) \neq \phi(\mathbf{D}_0)$, where \mathbf{S}^0 and \mathbf{S}_0 denote the closure and the interior of any \mathbf{S} , respectively, then those Borel subsets \mathbf{S} of \mathbf{X}^* which are discontinuity sets \mathbf{D} of any fixed ϕ

† The lengthy Jacobians occurring in the classical presentation of Jacobi's principle of the last multiplier are easily recognized to serve no other purpose than that of supplying an explicit analytical representation for the two-dimensional areal measures which are determined by the projection process.

are exceptional in the same sense as are the discontinuity points of a fixed monotone function of a single real variable.

One clearly obtains a distribution function $\phi = \phi(\mathbf{S}) = \phi_{uv}(\mathbf{S})$ if, on choosing any fixed path $x = x(t)$ in \mathbf{X}^* and two finite t -values, $t = u$ and $t = v (> u)$, one defines $\phi_{uv}(\mathbf{S})$ as the ratio of $\{u, v\}$ to $v - u$, where $\{u, v\}$ denotes, with reference to the given path $x = x(t)$, $-\infty < t < +\infty$, and to any Borel subset \mathbf{S} of \mathbf{X}^* , the measure of those points of the given t -interval of length $v - u$ for which the point $x(t)$ of the path is contained in \mathbf{S} . Since, by assumption, the path $x = x(t)$, $-\infty < t < +\infty$, lies in \mathbf{X}^* , the number $\phi_{uv}(\mathbf{S})$ represents the probability that a point of that portion of the path corresponding to $u < t < v$ should lie in the subset \mathbf{S} of \mathbf{X}^* . The given path is said to possess an asymptotic distribution function $\psi = \psi(\mathbf{S})$ if there exists a corresponding asymptotic probability. By this is meant the existence of a distribution function $\psi = \psi(\mathbf{S})$ on \mathbf{X}^* in such a way that for any fixed \mathbf{S} which is not a discontinuity set of ψ the value $\phi_{uv}(\mathbf{S})$ tends to the limit $\psi(\mathbf{S})$, as $v \rightarrow +\infty$, $-u \rightarrow +\infty$.

Needless to say, the asymptotic distribution function $\psi(\mathbf{S})$ of $x(t)$, $-\infty < t < +\infty$, (if it exists at all) depends, in general, on the choice of the path $x(t)$ or, what is the same thing, on the choice of that initial point x^0 in \mathbf{X}^* which determines $x(t)$ by the relation $x(t) = \tau^t x^0$ of §121; so that ψ will now be denoted by ψ_{x^0} .

Birkhoff's theorem states that, under the assumptions $\operatorname{div} f(x) \equiv 0$ and $0 < \mu(\mathbf{X}^*) < +\infty$ specified above, the asymptotic distribution function ψ_{x^0} exists for all those solution paths $x(t) = \tau^t x^0$ in \mathbf{X}^* for which the point x^0 of \mathbf{X}^* does not belong to a set of μ -measure 0; in other words, that "almost all" of the paths contained in the unrestricted invariant set \mathbf{X}^* possess an asymptotic distribution function.†

Incidentally, it is easy to see that if the Borel set \mathbf{S} is arbitrarily fixed, the function $\psi_x(\mathbf{S})$ of x is integrable (with respect to the ordi-

† In view of the extreme scarcity of "stable" motion in the usual sense (cf. §131 below), and also of the needs of statistical mechanics, it is natural to introduce, on the basis of the Ergodic Theorem, a notion of "distributional stability" of a solution, $x(t) = \tau^t x^0$, by the following requirement: The point x^0 does not belong to the excluded zero set and has the property that, if a variable point x of \mathbf{X}^* tends to x^0 in an arbitrary manner (provided that it avoids the excluded zero set), then $\psi_x(\mathbf{S})$ tends to $\psi_{x^0}(\mathbf{S})$ for every continuity set \mathbf{S} of the asymptotic distribution function ψ_{x^0} of $x(t) = \tau^t x^0$.

In order that this condition be satisfied for almost all x^0 , the metrical transitivity of τ^t (cf. §124 bis below) is sufficient but by no means necessary.

nary Lebesgue measure μ), and that its integral over the whole x -space \mathbf{X}^* has the value $\mu(\mathbf{S})/\mu(\mathbf{X}^*)$.

§123 bis. Another theorem which holds precisely under the assumption of the Ergodic Theorem is Poincaré's Recurrence Theorem. This theorem states, that, on the assumptions of §123, zero is the measure of the set of those points x^0 of \mathbf{X}^* for which the following condition is *not* satisfied: On placing $x(t) = \tau^t x^0$, one can find for every given date \bar{t} and for every $\epsilon > 0$ infinitely many dates $t_n = t_n(\bar{t}, \epsilon)$ which tend, as $n \rightarrow \pm \infty$, to $\pm \infty$ and are such that $|x(t_n) - x(\bar{t})| < \epsilon$ holds for every n .

While this Recurrence Theorem is obviously not implied by the wording of the Ergodic Theorem, it is a qualitative consequence of a quantitative fact (§124) which, when adjoined to the Ergodic Theorem, represents a refinement of the latter.

§124. In order to formulate this refinement, let Σ_{x^0} denote, for any x^0 not belonging to the zero set of the Ergodic Theorem, the set of those points \bar{x} of \mathbf{X}^* which have the property that any open set containing \bar{x} carries a positive asymptotic probability, i.e., the set of those \bar{x} for which $\psi_{x^0}(\mathbf{S}) > 0$ holds whenever \mathbf{S} is a sphere $|x - \bar{x}| < \rho$ about \bar{x} , where $\rho > 0$ is arbitrarily small but fixed. And let P_{x^0} denote, for any fixed x^0 , the closure of the path $x(t) = \tau^t x^0$, $-\infty < t < +\infty$, i.e., the set of those points x of \mathbf{X}^* which either are points, $x = \tau^t x^0$, of the path or are cluster points of such points. While it is obvious in itself that P_{x^0} contains Σ_{x^0} , it is not obvious from the Ergodic Theorem and the Recurrence Theorem together, that $\Sigma_{x^0} = P_{x^0}$ for almost all x^0 .

Nevertheless, $\Sigma_{x^0} = P_{x^0}$ is true for almost all x^0 .

This fact may be inferred from a careful perusal of Birkhoff's proof, though not from the usual wording, of the Ergodic Theorem, if use is made of the continuity properties of the transformation group τ^t (which are always assured by the conditions imposed on the differential equations defining the paths).†

† As a consequence of $\Sigma_{x^0} = P_{x^0}$, it is easy to infer from a general observation of Hadamard, that every subset Σ_{x^0} of the underlying closed, bounded unrestricted invariant set \mathbf{X}^* is an invariant set. Hence, it is clear from the definition of the set P_{x^0} , that if x^0, y^0 are any two points not belonging to the excluded zero set, then one of the two invariant sets $\Sigma_{x^0}, \Sigma_{y^0}$ must contain the other, if these two sets have at least one point in common.

Actually, it is possible that these two sets are always identical in case of at least one common point. In a terminology of Birkhoff, this possibility is ex-

§124 bis. It is natural to ask how can one characterize the particular case in which the asymptotic probability $\psi_{x^0}(\mathbf{S})$ is, for almost all x^0 and for every Borel set \mathbf{S} contained in \mathbf{X}^* , identical with the Euclidean volume measure $\mu(\mathbf{S})$ of \mathbf{S} (in view of the last remark of §123, this will be the case if and only if the asymptotic probability carried by an \mathbf{S} is independent of the initial condition x^0 for almost all x^0). It turns out that the answer is supplied by what is called the condition of metrical transitivity. This condition is defined by the requirement that the underlying \mathbf{X}^* should not contain any measurable invariant set $\tilde{\mathbf{X}}$ for which $\mu(\tilde{\mathbf{X}})$ is neither 0 nor the measure $\mu(\mathbf{X}^*)$ of the whole \mathbf{X}^* .

On the other hand, a path $x(t) = \tau^t x^0$ is called regionally transitive on \mathbf{X}^* if $P_{x^0} = \mathbf{X}^*$. And the system itself is called regionally transitive on \mathbf{X}^* if $P_{x^0} = \mathbf{X}^*$ holds for almost all x^0 contained in \mathbf{X}^* . According to §124, this is the case if and only if $\Sigma_{x^0} = \mathbf{X}^*$ for almost all x^0 contained in \mathbf{X}^* . This condition is obviously satisfied in the uniformly distributed case of metrical transitivity.

§125. The discussions of §126–§130 will be facilitated by first considering the example in which (1) is given, for $m = 4$, in terms of the partial derivatives $H_{x_i}(x)$ of the quadratic polynomial

$$(3) \quad H(x) \equiv H(x_1, x_2, x_3, x_4) = \frac{1}{2} \sum_{j=1}^2 (x_j^2 + \omega_j^2 x_{j+2}^2),$$

where $\omega_j = \text{const.} > 0$, in the form

$$(4) \quad x'_j = -H_{x_{j+2}}(x), \quad x'_{j+2} = H_{x_j}(x); \quad j = 1, 2, (\tfrac{1}{2}m = 2),$$

so that $x'_{j+2} = x_j$, $x''_{j+2} + \omega_j^2 x_{j+2} = 0$. Choosing \mathbf{X} to be the whole 4-dimensional Euclidean x -space, one sees from §89 that \mathbf{X} itself is an unrestricted invariant set \mathbf{X}^* in the sense of §120. The explicit form of the general solution $x(t) = x(x^0; t)$, where $x^0 = x(0)$, is easily found; whence the transition from (3), §79 to (5), §79 shows that $x' = f(x)$ has in the present case the $m = 4$ independent integrals

pressed by saying that, if x^0 does not belong to a zero set, then the path, $x(t) = \tau^t x^0$, is "minimal." According to a theorem of Birkhoff, this minimal property of a fixed path, $x(t) = \tau^t x^0$, can be characterized also directly, by the following property: There exists for every $\epsilon > 0$ an $l = l_\epsilon > 0$ in such a way that for any given t_0 one can find on any given t -interval of length l a point t at which $|x(t) - x(t_0)| < \epsilon$.

$$(5_j) \quad x_j \cos \omega_j t + x_{j+2} \omega_j \sin \omega_j t = x_j^0;$$

$$(5_{j+2}) \quad x_{j+2} \cos \omega_j t - x_j \omega_j^{-1} \sin \omega_j t = x_{j+2}^0,$$

where $j = 1, 2$. By the end of §82, elimination of t among (5₁)–(5₄) must lead to $m - 1 = 3$ independent conservative integrals $F_k(x) = c_k (= \text{const.})$.

First, elimination of t among (5_{*j*}) and (5_{*j+2*}) gives the pair of integrals $F_j(x) = c_j$, where $F_j = x_j^2 + \omega_j^2 x_{j+2}^2$; hence, $c_j \geq 0$ ($j = 1, 2$). Thus, the hypersurface $F_j(x) = c_j$ in \mathbf{X} is an hypercylinder, unless c_j vanishes. And the intersection of the two hypersurfaces $F_j(x) = c_j$ is a torus, if neither c_j vanishes. In any case, $F_j(x) = c_j$ is (the real portion of) an algebraic hypersurface for both $j = 1$ and $j = 2$. This holds no matter what are the values of the numerical constants $\omega_j > 0$ occurring in (3).

On the other hand, the structure of the remaining conservative integral, $F_3(x)$, depends very much on whether the ratio $\omega_1:\omega_2$ of the constant data ω_j of (4) is (i) a rational or (ii) an irrational number.

In case (i), let ω be the greatest common divisor of ω_1 and ω_2 ; so that $\omega_j = l_j \omega$, where l_1 and l_2 are relatively prime integers. Since the four functions $\sin \omega_j t$, $\cos \omega_j t$ are rational functions of $u = \tan \frac{1}{2} \omega t$, elimination of t between (5₁) and (5₂), say, leads to an integral $F_3(x)$ such that $F_3(x) = c_3$ is (the real part of) an algebraic hypersurface in \mathbf{X} . Roughly speaking, this hypersurface has the more self-intersections the higher is the commensurability $\omega_1:\omega_2$, i.e., the larger is $|l_1 - l_2|$. What is essential in what follows is not the algebraic character of this hypersurface but the fact that it has only a finite number of different “branches.”

In case (ii), there exists* for every $\epsilon > 0$ a pair of integers $l^{(j)} = l^{(j)}(\epsilon)$ such that $|\omega_1 l^{(1)} + \omega_2 l^{(2)}| > |l^{(1)}| > 1/\epsilon$. Hence, if one lets ϵ tend to zero,† the integral $F_3(x)$ which is obtained by elimination of t between (5₁) and (5₂) is easily shown to have the following property: There exists in the 4-dimensional x -space \mathbf{X} a domain such that, if x^0 is any point of this domain, then the (real and necessarily analytic) hypersurface $F_3(x) = c_3$ in \mathbf{X} , where $c_3 = F_3(x^0)$, has in every neighborhood of its point $x = x^0$ infinitely many different

* Cf. the footnote to §127 bis.

† And applies a straightforward argument familiar from similar applications of diophantine approximations; cf., e.g., the third footnote to §126.

“branches” which have, however, no manifolds of local ramification in the neighborhood of $x = x^0$. This peculiar situation will be cleared up in §126.

§126. It will be assumed in what follows that the region \mathbf{X} , on which the given single-valued m -vector function f of the position x is given as of class $C^{(1)}$, is an unrestricted invariant set \mathbf{X}^* of (1); cf. §120.

Let $F(x)$ be a single-valued scalar function of the position x on \mathbf{X} , and suppose that $F(x)$ is of class $C^{(1)}$ and nowhere constant on \mathbf{X} . A point x at which the gradient $F_x(x)$ vanishes is called a critical point of $F(x)$; these points, if any, are nowhere dense on \mathbf{X} . For any point x^0 of \mathbf{X} , let \mathbf{F}^{x^0} denote the “hypersurface” $F(x) = \text{const.}$ through x^0 ; more precisely, the set of *all* those points x of \mathbf{X} at which $F(x) = c$, where $c = F(x^0)$.

In particular, x^0 is an isolated point of \mathbf{F}^{x^0} if and only if the function $F(x)$ has at $x = x^0$ an isolated local extremum. If x^0 is an arbitrary critical point of $F(x)$, the topological structure of \mathbf{F}^{x^0} in the neighborhood of x^0 can be highly intricate.† If, on the other hand, x is not a critical point of $F(x)$, then the local existence theorem of implicit functions shows that \mathbf{F}^{x^0} is in the vicinity of x^0 a single connected piece (“branch”) of an $(m - 1)$ -dimensional surface (i.e., hypersurface), with a definite normal and with no self-intersections.

How is it, then, possible that the integral $F_3(x)$ of (4) is, in the case (ii) considered at the end of §125, such that the corresponding \mathbf{F}^{x^0} has in any vicinity of x^0 infinitely many different “branches,” at least if one chooses x^0 in a certain x^0 -domain? (The situation seems to be a paradox, since, \mathbf{F}^{x^0} being obtained by elimination of t between the analytic relations (5₁)–(5₂), reasons of analyticity insure that the x^0 -domain in question can be so chosen as to contain no critical points of the function $F_3(x)$, which is regular at x^0 .) The answer is implied by the warning given at the end of §82, the situation being as follows:

Let $x = x(t)$ be the solution path through $x^0 = x(0)$, and suppose that x^0 is not an equilibrium point of (1). Then $x(t^{(1)}) = x(t^{(2)})$ is

† Notice that \mathbf{F}^{x^0} can be a nowhere dense perfect set even when $F(x)$ is of class $C^{(\infty)}$ (i.e., of class $C^{(\nu)}$ for every ν). If $F(x)$ is regular analytic on \mathbf{X} and x^0 is a critical point of $F(x)$, the dimension number of \mathbf{F}^{x^0} close to x^0 can be any integer (≥ 0) less than m ; while the situation in the case of several critical points is, in the large, strongly restricted by Morse’s well-known index relations.

impossible for $t^{(1)} \neq t^{(2)}$. On the other hand, it is quite possible that there exist two sequences of dates t_n^I, t_n^{II} which tend with n to ∞ and are such that $|x(t) - x^0| < 1/n$ or $|x(t) - x^0| > \text{const.} > 0$ according as $t_n^I \leq t \leq t_{n+1}^I$ or $t_n^{II} \leq t \leq t_{n+1}^{II}$ (nothing is said as to t not contained in one of these intervals). Now, if one applies local existence theorems in the t -neighborhood of every $t^{(n)} = \frac{1}{2}(t_n^I + t_{n+1}^I)$ and in the x -neighborhood of every $x^{(n)} = x(t^{(n)})$, then, since $|x^{(n)} - x^0| \rightarrow 0$, nothing hinders the clustering* of *different* branches of F^{x^0} at x^0 . This the more as the elimination of t in the neighborhood of the $\frac{1}{2}(t_n^{II} + t_{n+1}^{II})$ might lead to distant critical points (or even to singularities) of $F(x)$ which correspond to distant ramifications of F^{x^0} ; while the branches arising from these distant ramifications,† when continued along the solution path, can easily reach the points $x^{(n)}$ which correspond to the $\frac{1}{2}(t_n^I + t_{n+1}^I)$ and cluster at x^0 .

§127. The example of §125 is simple enough to make one think that this situation is not a “degenerate” but rather the “general” case, when $f(x)$ in (1) is unspecified. A conjecture to this effect, though a central conviction of modern dynamics, has escaped all efforts thus far made to provide a satisfactory proof. The conviction in question is that (in view of the postulates of classical statistical mechanics but first of all in view of the investigations of Poincaré, Hadamard, Levi-Civita and Birkhoff, as well as in view of detailed investigations concerning Fuchsian groups or geodesics on surfaces of negative curvature) regional, if not metrical, transitivity (§124 bis) characterizes a “generic” system (in this connection, cf. §131).

§127 bis. Consider, for instance, the example (2), §121 bis on the assumption that \mathbf{X} is thought of as the torus, $0 \leq x_1 < 1, \dots, 0 \leq x_m < 1$, obtained by reduction modulo $(\pi_1, \dots, \pi_m) = (1, \dots, 1)$, and suppose that $s = m$, where the non-negative integer $s (\leq m)$ is defined by the property that, with reference to the rational field,

* The situation may well be compared to that arising in case of the inverse function $z = z(w)$ of a transcendental entire function $w = w(z)$ which has the property that, while a certain $w = w^0$ is a regular point of $z(w)$ on every point of the Riemann surface, the z -values attained by $z(w)$ in the neighborhoods of w^0 on the different sheets form z -domains which cluster at the point $z = z(w^0)$ of the z -plane.

† The situation might be compared with the one arising in case of unramified Abelian integrals which, though locally uniformized by the Riemann surface of the underlying algebraic function, are not single-valued functions of the position on this Riemann surface. The actual situation is, however, closer to that arising in connection with the hyperelliptic inversion problem of Jacobi; cf., in fact, the footnote to §128.

there are exactly s linearly independent numbers among the λ_i which constitute the components of the m -vector $f(x) = \lambda = \text{const.}$ Then every solution path is a loxodrome which is* regionally transitive on \mathbf{X} .

If, in particular, $m = 2$ and the equation of a solution path on the (x_1, x_2) -torus is written in the form $F(x_1, x_2) = c$, one can think of $F(x_1, x_2)$ as a conservative integral obtained by the elimination process described at the end of §82. Since $\lambda_1:\lambda_2$ is irrational, the transitive path $F(x_1, x_2) = c$ does not intersect itself on the torus.

§128. Using the assumptions and notations of §126, one sees that if $m = 2$, then \mathbf{F}^{x^0} is the solution path through x^0 , and that if $m > 2$, then this solution path can, at least locally, be thought of as the common part of the $m - 1$ sets \mathbf{F}^{x^0} which belong to $m - 1$ conservative independent integrals $F(x)$; cf. §82. Correspondingly, an integral $F(x)$ is valuable only insofar as it can enable one to make predictions concerning the possible future (or past) positions of the points $x = x(t)$ of the solution path which goes at $t=0$ through x^0 (the case $x(t) \equiv x^0$ of an equilibrium solution being not excluded). From this point of view, the knowledge of an integral of the type of $F_3(x)$ in case (ii), §125, or of $F(x_1, x_2)$ at the end of §127 bis, is quite worthless.

Those integrals of (1) which are not worthless in this sense will be called isolating.† A detailed and explicit definition of an isolating integral would presuppose a topology in the large for the underlying unrestricted invariant sets. Actually, the whole question is of true significance only under restrictions of analyticity.

In order that an integral $F(x)$ of (1) be isolating, it is neither necessary nor sufficient that the function $F(x)$ be a single-valued function

* In view of what is called Kronecker's approximation theorem (for $m = 2$ one obtains a standard property of continued fractions; cf. the inequalities for $l^{(1)}, l^{(2)}$, used in the case (ii) of §125).

Actually, the transitivity is, in the present case, not only regional (Kronecker) but metrical as well (Weyl); cf. §124 bis. It is known (H. Bohr) that in the present case the regional transitivity implies the metrical transitivity as an immediate consequence. Also notice that the zero sets excluded in §123–§124 are vacuous in the present case.

† Unfortunately, the adjective used in the existing literature is "uniform" = "eindeutig," i.e. "single-valued"; an adjective which describes the actual situation less correctly than does "isolating," and is responsible for frequent misunderstanding by theoretical physicists of the results of Poincaré concerning "intégrales uniformes." Actually, Poincaré's "uniforme" is patterned after a time-honored terminology of Jacobi concerning the inversion of elliptic and hyperelliptic integrals, respectively.

of the position on \mathbf{X} , i.e., that \mathbf{F}^{x^0} be free of self-intersections for every x^0 . This is shown, respectively, by the isolating example $F_3(x)$ in the case (i) of §125 and by the non-isolating example $F(x_1, x_2)$ at the end of §127 bis.

§129. Classical researches have succeeded in establishing certain negative results of a type which can be illustrated by the simplest of the theorems of Bruns, subsequently refined by Painlevé. This simplest of the theorems in question states that that system $x' = f(x)$ which represents the problem of more than two bodies (in terms of Cartesian coordinates) does not possess conservative algebraic integrals $F(x)$ distinct from the algebraic consequences of those, seven in number, which were known by the middle of the eighteenth century, at least. It must, however, be said that the elegant negative results of this arithmetical type do not have any dynamical significance. For all that is of dynamical interest is an enumeration of all those independent integrals $F(x)$ which are isolating. Now, even if $f(x)$ is algebraic, the algebraic character of an integral $F(x)$ of (1), though sufficient, is by no means necessary for an $F(x)$ which is an isolating integral.

§130. If the system (1) of m scalar differential equations has l , but does not have $l + 1$, isolating integrals $F(x)$, and if one excludes the trivial case $f(x) \equiv 0$ in which the number of all conservative independent integrals is m instead of being, as in every other case, $m - 1$ (cf. the end of §82), then the system (1) is called $(m - 1 - l)$ -fold primitive or, equivalently, l -fold imprimitive; correspondingly, (1) is called primitive if $l = 0$. The ideal case, where all $m - 1$ independent local integrals $F(x)$ happen to be significant in the large, is the case of $(m - 1)$ -fold imprimitivity; while $l = 0$ clearly is a necessary (and, as far as present knowledge goes, possibly sufficient) condition for the existence of paths regionally transitive in \mathbf{X} (cf. §127).

In the torus example of §127 bis, one has $l = m - s$; so that the system is primitive in case the λ_i are linearly independent ($s = m$). In the example of §125, where $m = 4$, one has $l = 3$ and $l = 2$ in the respective cases (i) and (ii); so that (3), (4) define a 0-fold or 1-fold primitive system (1) according as $\omega_1:\omega_2$ is rational or irrational.*

* A statistical approach to dynamical systems of a given degree of imprimitivity is developed by Levi-Civita's theory of what Ehrenfest has introduced as adiabatic invariants.

Points of Stability

§131. There are about a dozen different definitions of “stability,” which are all useful but have little, if anything, to do with one another; every definition requiring a different desirable property either of a solution or of a collection of solutions.

One of the oldest definitions of stability of a given solution $x = \bar{x}(t)$ of $x' = f(x)$ is obtained by requiring that the facts which at the end of §84 were seen to hold for a fixed large t -interval should hold for $-\infty < t < +\infty$. In other words, a given solution $x = \bar{x}(t)$ of $x' = f(x)$ is called stable in this sense, if it has the following properties: There exists for every $\epsilon > 0$ a $\delta = \delta_\epsilon > 0$ such that if $x = x(t)$ is any solution of $x' = f(x)$ for which the initial position $x(0)$ satisfies, with reference to the initial position $\bar{x}(0)$ of the given solution $x = \bar{x}(t)$, the inequality $|x(0) - \bar{x}(0)| < \delta$, then (i): $x = x(t)$ is an unrestricted solution in the sense of §119, and (ii): one has $|x(t) - \bar{x}(t)| < \epsilon$ for $-\infty < t < +\infty$. (Choosing $t = 0$, one sees that $\delta \leq \epsilon$; choosing $x(0) = \bar{x}(0)$, one sees that $x = \bar{x}(t)$ itself must be unrestricted.)

This definition of stability seems to be the most natural one. Actually, it is not natural at all, since it requires too much. In fact, everything that is known from Poincaré’s geometrical theory of real differential equations and from the parallel, though more difficult, theory of surface transformations (Poincaré, Hadamard, Levi-Civita, Birkhoff) points in the direction that condition (ii) cannot be satisfied except in highly exceptional cases. Even in the restricted problem of three bodies, not a single solution is known to be stable.

The situation seems to be that, precisely in the interesting cases, condition (ii) becomes violated for Diophantine reasons; reasons which depend on properties of irrational numbers and appear immediately on introduction of angular variables. This remark is illustrated by the fact that the only useful criterion which is known to be sufficient for (i)–(ii) concerns merely the case in which the given solution $x = \bar{x}(t)$ of $x' = f(x)$ is an equilibrium solution in the sense of §83.

§132. In order to formulate this criterion, let $\Sigma_1, \Sigma_2, \dots$ be a sequence of sets in the x -space, with the property that a given point x^0 is an interior point* of every Σ_n , while Σ_n shrinks,† as $n \rightarrow \infty$, to

* By this is meant that if $S(\eta)$ denotes the sphere $|x - x^0| < \eta$ about x^0 , then every point x of $S(\eta_n)$ is contained in Σ_n , if $\eta_n > 0$ is sufficiently small.

† By this is meant that if $S(\eta)$ is defined as in the preceding footnote, then

the point x^0 . Suppose that the given point x^0 represents an equilibrium solution $x(t) \equiv x^0$ of $x' = f(x)$, i.e., that $0 = f(x^0)$. Suppose further that every Σ_n is an invariant set of $x' = f(x)$ in the sense of §81. Then the solution $x(t) \equiv x^0$ of $x' = f(x)$ is stable in the sense of §131. This becomes clear by comparing the definition of an invariant set with the last remark of §120.

§133. The sufficient condition of §132 for the stability of $x(t) \equiv x^0$ is necessary as well. In other words, if the equilibrium solution $x(t) \equiv x^0$ of $x' = f(x)$ is stable, then there exists a sequence of invariant sets Σ_n which are domains and shrink, as $n \rightarrow \infty$, to the invariant point x^0 .

In order to see this, let $\mathbf{S}'(\eta)$ denote, for a fixed sufficiently small $\eta > 0$ and for any fixed t , the set of those points of the x -space for which $x' = f(x)$ has a solution path passing at the given t and at $t = 0$ through some point of $\mathbf{S}'(\eta)$ and through some point of $\mathbf{S}(\eta)$, respectively, where $\mathbf{S}(\eta)$ denotes the sphere $|x - x^0| < \eta$. Let $\mathbf{R}(\eta)$ be the set of those points of the x -space which are contained in at least one $\mathbf{S}'(\eta)$, where η is fixed and t varies from $-\infty$ to $+\infty$. Thus, $\mathbf{R}(\eta)$ is a collection of unrestricted paths (namely, of those paths each of which has a point within $\mathbf{S}(\eta)$ at a suitable t). Consequently, $\mathbf{R}(\eta)$ is an invariant set. Furthermore, $\mathbf{R}(\eta)$ shrinks, as $\eta \rightarrow +0$, to the invariant point x^0 , since $x(t) \equiv x^0$ is supposed to be a stable solution of equilibrium. Finally, x^0 is an interior point of $\mathbf{R}(\eta)$, since $\mathbf{R}(\eta)$ contains the sphere $\mathbf{S}(\eta)$. Accordingly, the sets Σ_n whose existence has to be proved may be obtained by placing, for instance, $\Sigma_n = \mathbf{R}(n^{-1})$ for every sufficiently large n .

§134. It will now be shown that if $x(t) \equiv x^0$ is an equilibrium solution of $x' = f(x)$, and if $x' = f(x)$ has a conservative integral $F(x) = \text{const.}$ such that the function $F(x)$ of the position in the x -space has at the point $x = x^0$ either an isolated maximum or an isolated minimum, then the solution $x(t) \equiv x^0$ is stable* in the sense of §131.

there exists for every $\eta > 0$ an integer N_η such that $\mathbf{S}(\eta)$ contains Σ_n for every $n > N_\eta$.

* A consequence of this theorem is that the solar system is stable, if only the "secular" perturbations are taken into account. (On the assumption that only the linear secular perturbations are considered, the stability of the solar system was known to Lagrange [and Laplace]. The observation that, due to the above theorem of Minding [and Dirichlet], all non-linear secular perturbations may be included, was made by Bruns.)

In order to prove this, only the case of a minimum needs to be considered, since $F(x)$ may be replaced by $-F(x)$. Furthermore, it may be assumed that $F(x^0) = 0$, since $F(x)$ may be replaced by $F(x) + \text{constant}$. Thus, $F(x) > 0$ for every x sufficiently close to, but distinct from, x^0 . Since $F(x)$ is a continuous function of the position in the x -space, it follows that there exists for every sufficiently small $\zeta > 0$ a domain $\Sigma = \Sigma(\zeta)$ which contains a vicinity of the point $x = x^0$, shrinks to this point as $\zeta \rightarrow 0$, and is such that $F(x) < \zeta$ or $F(x) = \zeta$ according as x lies in $\Sigma(\zeta)$ or on the boundary of $\Sigma(\zeta)$. Since $F(x) = \text{const.}$ is an integral of $x' = f(x)$, it follows from §80–§82 that $\Sigma(\zeta)$ is an invariant set of $x' = f(x)$ for every fixed small $\zeta > 0$. This completes the proof, since it is now clear that the conditions imposed on the Σ_n in §132 are satisfied by choosing $\Sigma_n = \Sigma(\zeta_n)$ and $\zeta_n = n^{-1}$, say.

§134 bis. Suppose, for instance, that $x' = f(x)$ is represented by (1), §91, where $H_t \equiv 0$, and that $H(p, q) = T - U$, where T is a positive definite quadratic form in the components of $p = (p_1, \dots, p_n)$, while U is a function of $q = (q_1, \dots, q_n)$ having at $q = (0, \dots, 0)$ an isolated maximum. Then the integral $F(x) = \text{const.}$ represented by (3), §92 obviously has at $x = 0$ an isolated minimum, and so §134 is applicable to $F \equiv H$ at $x^0 = 0$.

Notice, however, that §134 might become applicable to a conservative canonical system at some $x(t) \equiv x^0$ also when the condition of §134 is not satisfied by the energy integral (3), §92, but is satisfied by another* integral $F(x) = \text{const.}$

§135. It should be mentioned that the sufficient condition of §134 for the stability of $x(t) \equiv x^0$ is not a necessary condition. For let $x' = f(x)$ be given as a conservative Hamiltonian system with a single degree of freedom, having the Hamiltonian function $H(x) = H(p, q) = \frac{1}{2}p^2 - U(q)$, where $U(q) = \exp(-q^{-2}) \cos(q^{-1})$ for $q \neq 0$ and $U(0) = 0$ (so that all derivatives of $U(q)$ exist for every q). Then $p \equiv 0, q \equiv 0$ is an equilibrium solution, since $H_p(0, 0) = 0, H_q(0, 0) = 0$. It is a stable solution, since, on cutting the energy surface $H = H(p, q)$ in a Cartesian (p, q, H) -space by a suitable sequence $H = h_n (= \text{const.})$ of planes, one readily verifies that the

* An important instance to this effect is the application mentioned in the footnote to §134, where $F(x) = \text{const.}$ must be chosen so as to correspond not to the energy integral but to the conservation integral of angular momentum.

condition of §132 is satisfied. In fact, one can choose a sequence of energy constants h_n which tend, as $n \rightarrow \infty$, to the energy constant $h = 0$ of $p \equiv 0, q \equiv 0$ and are such that the "curve" $H(p, q) = h_n$ in the (p, q) -plane has a closed branch surrounding a domain Σ_n which, in turn, contains the point $(p, q) = (0, 0)$ and tends, as $n \rightarrow \infty$, to this point. Nevertheless, the definition of $U(q)$ shows that the function $H = \frac{1}{2}p^2 - U(q)$, which is the integral $F(x)$ of §134, has neither a maximum nor a minimum at $(p, q) = (0, 0)$; while there cannot exist a conservative integral independent of the energy integral $H(p, q) = \text{const.}$, since the latter is the equation of the solution in the (p, q) -plane.

§135 bis. It is easy to see that, in the example of §135, the stable equilibrium point is a cluster point both of unstable and of stable equilibrium points. Actually, it is not known whether or not the sufficient condition of §134 is necessary as well in case there is no clustering of equilibria (e.g., in case the system is regular analytic). Cf. also §477 bis below.

§136. Suppose that $x' = f(x)$ has the equilibrium solution $x(t) \equiv x^0$, and let A be the constant m -matrix which represents the Jacobian matrix of the m -vector $f(x)$ with respect to x at the point $x = x^0$. Then the Jacobi equations are $\xi' = A\xi$, by §89. Hence, one might expect, by §85, that the equilibrium solution $x(t) \equiv x^0$ of $x' = f(x)$ is stable in the sense of §131 whenever all characteristic exponents of $\xi' = A\xi$ are of the stable type in the sense of §89 and A does not have multiple elementary divisors. In fact, this pair of conditions for A clearly is necessary and sufficient for the boundedness of every solution $\xi = \xi(t)$, $-\infty < t < +\infty$, of $\xi' = A\xi$; so that this pair of conditions for the constant matrix A seems to be sufficient for the stability (in the sense of §131) of the equilibrium solution $x(t) \equiv x_0$ of $x' = f(x)$.

However, simple examples show that the theorem in question is false. An example to this effect may indeed be so chosen that $x' = f(x)$ is a canonical system.

§136 bis. To this end, let $x' = f(x)$ be given as the conservative system

$$(1_1) \quad x'_j = -H_{x_{j+2}}, \quad x'_{j+2} = H_{x_j};$$

$$(1_2) \quad H = \frac{1}{2}(x_1^2 + x_3^2) - (x_2^2 + x_4^2) + \frac{1}{2}(x_4x_3^2 - x_4x_1^2 - 2x_1x_2x_3)$$

with $n = \frac{1}{2}m = 2$ degrees of freedom; ($j = 1, 2$). Since all four partial derivatives H_{x_i} of (1₂) are seen to vanish at the origin, $x(t) \equiv 0$ is an equilibrium solution of (1₁). According to §101, the corresponding Jacobi equations are obtained by replacing x in (1₁) by ξ , and H by the quadratic part of the cubic polynomial (1₂), i.e., by $\frac{1}{2}(\xi_1^2 + \xi_3^2) - (\xi_2^2 + \xi_4^2)$. Hence, the explicit form of the Jacobi equations $\xi' = A\xi$ is

$$(2) \quad \xi'_1 = -\xi_3, \quad \xi'_2 = 2\xi_4, \quad \xi'_3 = \xi_1, \quad \xi'_4 = -2\xi_2.$$

It is seen from (2) that the $m = 4$ characteristic numbers of A are $s = \pm \sqrt{-1}$ and $s = \pm 2\sqrt{-1}$, hence all distinct and of the stable type (§89). Nevertheless, the equilibrium solution $x(t) \equiv 0$ of (1₁) is not stable in the sense of §131.

In fact, on calculating the four partial derivatives $H_{x_i}(x)$ of the cubic polynomial (1₂), one easily verifies that (1₁) admits the particular solution $x = x(t)$ given by

$$\begin{aligned} x_1(t) &= 2^{\frac{1}{2}}t^{-1} \cos t, & x_2(t) &= -t^{-1} \cos 2t, \\ x_3(t) &= 2^{\frac{1}{2}}t^{-1} \sin t, & x_4(t) &= t^{-1} \sin 2t. \end{aligned}$$

This solution of (1₁) tends, as $t \rightarrow \pm \infty$, to the equilibrium solution $x(t) \equiv 0$, while $x_1(t)$, $x_2(t)$ become infinite as $t \rightarrow \pm 0$. But the system is conservative; so that $x(t)$ may be replaced by $x(t - t_0)$ where t_0 is an arbitrary constant. Hence, on choosing t_0 large, one sees that neither of the conditions (i), (ii) of §131 is satisfied by $x(t) \equiv 0$.

Characteristic Exponents

§137. The complications mentioned at the end of §79 cannot arise in case of a linear system $\xi' = A(t)\xi$, where $\xi = \xi(t)$ is an unknown m -vector and $A(t)$ a given m -matrix which is supposed to be continuous for $0 \leq t \leq t^*$ (or $t^* \leq t \leq 0$), say. In fact, no matter what is the initial condition $\xi(0)$, the corresponding solution $\xi(t)$ exists and is unique, for all t between $t = 0$ and $t = t^*$.

Actually, $\xi(t)$ may be obtained from $\xi(0)$ by a linear transformation, given by an m -matrix $R(t)$ which is independent of $\xi(0)$ and has a determinant which is expressible in terms of the trace† of A :

$$(1_1) \quad \xi' = A(t)\xi; \qquad (1_2) \quad \xi(t) = R(t)\xi(0);$$

† The trace of an m -matrix $B = (b_{ik})$ is defined by $\text{tr } B = b_{11} + b_{22} + \cdots + b_{mm}$.

$$(1_3) \quad \det R(t) = \exp \int_0^t \operatorname{tr} A(\bar{t}) d\bar{t}.$$

In fact, application of the method of successive approximations to (1₁) gives, for all t between 0 and t^* ,

$$(2_1) \quad R(t) = \sum_{k=0}^{\infty} D_k(t);$$

$$(2_2) \quad D_{k+1}(t) = \int_0^t A(\bar{t}) D_k(\bar{t}) d\bar{t}, \quad D_0(t) = E,$$

where E is the unit matrix. And the m -matrix series (2₁), defined by the recursion formula (2₂), and the derived series $R'(t) = \sum D'_k(t)$ have for all t between 0 and t^* a convergent exponential majorant series which is independent of t . Furthermore, substitution of (1₂), (2₁) into (1₁) gives the identity $R'(t) = A(t)R(t)$; hence, differentiation of $\det R(t)$ shows that $(\det R)' = (\det R)(\operatorname{tr} A)$. This proves (1₃), since $R(0) = E$, by (2₁)–(2₂).

§138. An m -matrix $X(t)$ whose columns are constituted by m linearly independent solutions $\xi(t)$ of (1₁) is called a fundamental matrix of (1₁). Since this is the case if and only if $X'(t) = A(t)X(t)$ and $\det X(t) \neq 0$, it is clear that another m -matrix, $Z(t)$, is a fundamental matrix of (1₁) if and only if there exists a constant m -matrix C such that $Z(t) = X(t)C$ and $\det C \neq 0$ (principle of superposition). Since $R'(t) = A(t)R(t)$ by §137, and since (1₃) cannot vanish, $R(t)$ is a fundamental matrix. Hence, the definition of a fundamental matrix $X(t)$ of (1₁) may be written in either of the equivalent forms

$$(3_1) \quad X'(t) = A(t)X(t), \quad \det X(t) \neq 0;$$

$$(3_2) \quad X(t) = R(t)C, \quad \det C \neq 0, \quad (R(0) = E),$$

where C is a non-singular constant matrix which is uniquely determined by $X(t)$. In fact, $C = R^{-1}X$, since $\det R \neq 0$, by (1₃). It is also seen from (3₂) and (1₃) that $\det X(t) \neq 0$ for every t ; so that m solutions $\xi(t)$ of (1₁) cannot be linearly independent for a single t unless they are linearly independent for every t .

§139. Let $X(t)$ be a fundamental matrix of (1₁), and C any non-singular constant matrix. Right-hand multiplication of $X(t)$ by C means, by §138, transition to another fundamental matrix, $Z(t) = X(t)C$, of (1₁). On the other hand, left-hand multiplication of

$X(t)$ by C , i.e., transition from $X(t)$ to $Y(t) = CX(t)$, means transition from the system (1_1) to another system, in which the coefficient matrix $A(t)$ is replaced by $CA(t)C^{-1}$. In fact, (3_1) may be written in the form

$$(4_1) \quad Y'(t) = B(t)Y(t); \quad (4_2) \quad B(t) = CA(t)C^{-1}; \quad (4_3) \quad Y(t) = CX(t).$$

It is clear that the above considerations remain valid also when $A(t)$, C are allowed to be complex. While $A(t)$ will always be supposed to be real, it will, in §144, be convenient to allow complex C , if a certain real matrix, which will be defined in terms of $A(t)$, happens to have complex characteristic numbers.

§140. Suppose that the continuous coefficient matrix $A(t)$ of (1_1) is given for $-\infty < t < +\infty$ as periodic, say $A(t) = A(t + \tau)$. The period $\tau \neq 0$, which is then not uniquely determined,* will be supposed to have a fixed value. Since (3_1) remains valid if one writes $t + \tau$ for t , it is clear from $A(t + \tau) = A(t)$ that $X(t + \tau)$ is, for every fundamental matrix $X(t)$ of (1_1) , a fundamental matrix of (1_1) . It follows, therefore, from §138 that there exists for every fundamental matrix $X(t)$ of (1_1) a unique non-singular matrix $\Gamma = \Gamma_X$ such that the relation

$$(5) \quad X(t + \tau) = X(t)\Gamma_X, \quad \text{where} \quad \det \Gamma_X \neq 0, \quad \Gamma_X = \text{const.},$$

is an identity in t . This unique Γ_X is called the monodromy matrix of the fundamental matrix $X(t)$ (with reference to the given period τ of A). In particular,

$$(6) \quad \Gamma_R = R(\tau), \quad \text{since} \quad R(0) = E, \quad \text{by} \quad (1_2).$$

§141. According to §138, the most general fundamental matrix of (1_1) is $X(t)C$, where $C = \text{const.}$ and $\det C \neq 0$. Furthermore, the monodromy matrix Γ_{XC} of $X(t)C$ is

$$(7) \quad \Gamma_{XC} = C^{-1}\Gamma_X C,$$

by (5). Hence, if a constant matrix Γ is a monodromy matrix of some fundamental matrix of (1_1) , another constant matrix is the monodromy matrix of a suitable fundamental matrix of (1_1) if and only if it is of the form $C\Gamma C^{-1}$ for some constant non-singular C , i.e., if and only if it has the same characteristic numbers and elementary divisors (invariant factors) as Γ . Correspondingly, these charac-

* In fact, any multiple of τ is again a period.

teristic numbers (with their proper multiplicities) and elementary divisors are called the invariants of the monodromy group of (1_1) , this group being defined by (7) with reference to the fixed period τ of $A(t)$; cf. (5).

§142. In particular, the m characteristic numbers of Γ_X (which are independent of the choice of $X(t)$) are called the multipliers of (1_1) with reference to τ . None of these multipliers vanishes, since their product is $\det \Gamma$ in view of their definition $\det (sE - \Gamma) = 0$, where $\det \Gamma \neq 0$, by (5), (7).

Since the multipliers of (1_1) can be defined as the characteristic numbers of the matrix (6), which is real by (2_1) – (2_2) , it is clear that complex multipliers can occur only in conjugate pairs; and a similar remark holds for the elementary divisors which belong to complex multipliers, if any.

§143. Denoting by s_j , where $j = 1, \dots, m$, the multipliers of (1_1) , and using the fact that every $s_j \neq 0$, one can introduce m numbers λ_j by placing, with reference to the fixed value of the period $\tau \neq 0$ of $A(t)$,

$$(8) \quad \lambda_j = \tau^{-1} \log s_j, \quad \text{so that} \quad s_j = e^{\lambda_j \tau} (\neq 0); \quad j = 1, \dots, m.$$

It is understood that the m numbers λ_j , which are called the characteristic exponents of (1_1) , are determined only mod $2\pi i/\tau$. Correspondingly, by $\lambda_j = \lambda_k$ will be meant that $\lambda_j - \lambda_k$ is a real integral multiple of $2\pi i/\tau$. For instance, if j is fixed, then $s_j = 1$ if and only if $\lambda_j = 0$ (or $\lambda_j = 2\pi i/\tau$); while $s_j = -1$ if and only if $\lambda_j = \pi i/\tau$.

If $|s_j| = 1$ for a fixed j , then, whether s_j is complex or real, s_j or λ_j will be called of the stable type. According to (8), this is the case if and only if λ_j is purely imaginary, including 0. It is also clear from (8) that if λ_j is real and $\tau > 0$, then $\lambda_j \gtrless 0$ according as $s_j \gtrless 1$, where $s_j > 0$. Hence, if no determination of λ_j can be chosen as real or as purely imaginary, then either $s_j < 0$ or $s_j = a_j + ib_j$, where $a_j \neq 0$, $b_j \neq 0$ are real. Actually, it is seen from (8) that $s_j < 0$ if and only if the imaginary part of λ_j is $\pi i/\tau$.

§144. Since Γ_X can be replaced by any matrix (7), one may assume the fundamental matrix $X(t)$ of (1_1) so chosen that its monodromy matrix Γ_X has the Jordan normal form. Then the diagonal elements of Γ_X are the multipliers s_1, \dots, s_m , and the line parallel to, and bordering from above, the diagonal of Γ_X contains only the numbers 0 and 1 (possibly only 0's or only 1's); while the elements of Γ_X which

are not contained in these two parallel lines are all 0. Let s be one of the s_j , and let the multiplicity of this s be denoted by l (≥ 1); so that the first l diagonal elements of Γ_X may be assumed to be equal to s . Let s belong to distinct elementary divisors of respective multiplicities h_1, \dots, h_d ; so that $h_1 + \dots + h_d = l$, where $l \geq 1$, $d \geq 1$, and every $h \geq 1$. Consider in the matrix Γ_X , which has the Jordan normal form, one of the blocks belonging to a fixed h . It may be assumed that this block is the first block of Γ_X , i.e., the h_1 -th section of Γ_X . Then, on denoting by $\xi_1(t), \dots, \xi_m(t)$ the m -vectors which constitute the successive columns of the m -matrix $X(t)$, one sees from (5) that

$$(9_1) \quad \xi_1(t + \tau) = s\xi_1(t);$$

$$(9_2) \quad \xi_g(t + \tau) = \xi_{g-1}(t) + s\xi_g(t); \quad g = 2, \dots, h_1,$$

where (9₂) is missing in the case $h_1 = 1$ of a simple elementary divisor. Now, it is easily verified from (8) that (9₁)–(9₂) are equivalent to

$$(10_1) \quad \xi_1(t) = e^{\lambda t} \phi_{11}(t);$$

$$(10_2) \quad \xi_g(t) = e^{\lambda t} \sum_{k=1}^g t^{k-1} \phi_{gk}(t); \quad g = 2, \dots, h_1,$$

where λ is the characteristic exponent belonging to the multiplier s , the $\phi(t)$ are certain m -vector functions of t which have the common period τ , and $\phi_{gg}(t) \not\equiv 0$ for $g = 1, \dots, h_1$. On proceeding in this manner, first for each of the remaining $d - 1$ (≥ 0) blocks which belong to the fixed characteristic number s of Γ_X , and then for each of the distinct values among the s_j , one clearly arrives at the following results:

§144 bis. A number $s = e^{\lambda\tau}$ is multiplier of (1₁) if and only if (1₁) has a solution of the form $\xi(t) = e^{\lambda t} \phi(t)$, where $\phi(t)$ has the same period τ as $A(t)$, and $\phi(t) \not\equiv 0$. The general solution of (1₁) is a linear combination of m linearly independent solutions of the form $e^{\lambda t} \phi(t)$ if and only if the elementary divisors of the monodromy group are all simple. If at least one of the elementary divisors is multiple, the general solution of (1₁) contains “secular” terms, i.e., terms which contain, besides periodic or exponential functions of t , rational polynomials of t ; the exponent of the highest power of t being exactly $h - 1$, if h is the multiplicity (that is, if $h - 1$ is the degree) of the corresponding elementary divisor.

§145. Suppose that $A(t)$ is independent of t . Then (1_1) , (2_1) , (1_2) reduce to*

$$(11_1) \xi' = A\xi, (A = \text{const.}); (11_2) R(t) = e^{tA}; (11_3) \xi(t) = e^{tA}\xi(0).$$

Since the assumption $A(t + \tau) = A(t)$ of §140 is satisfied by $A = \text{const.}$ for every τ , and, since the $\phi(t)$ of (10_1) – (10_2) are of period τ , every $\phi(t) = \text{const.}$ The characteristic exponents λ , which are by (10_1) – (10_2) determined only modulo $2\pi i/\tau$ (cf. §143), become uniquely determined, since this modulus is arbitrary. Actually, the λ are the characteristic numbers of A ; cf. §89. On the other hand, the monodromy group, hence also the set of the multipliers s , becomes completely undetermined, since τ in (5) is now arbitrary.

§146. Let $x = x(t)$ be a given solution of a system $x' = f(x)$. The corresponding Jacobi system (8), §85, defined by (9), §85 and $x(\bar{x}^0; t) = x(t)$, may be identified with (1_1) , §137.

Thus, (10), §85 shows that the particular fundamental matrix (2_1) , §137 becomes identical with the matrix (7), §85. This fact is, in view of the result (6), §85, fundamental in the applications.

§147. Suppose, in particular, that the given solution $x = x(t)$ of $x' = f(x)$ is periodic, $x(t + \tau) = x(t)$. Then the assumption $A(t + \tau) = A(t)$ of §140 is satisfied,† and so one can speak of the characteristic exponents $\lambda_1, \dots, \lambda_m$ of the periodic solution $x(t)$ of $x' = f(x)$, the λ being referred to a fixed period τ of $A(t)$.

These characteristic exponents λ , or the corresponding multipliers $s = e^{\lambda\tau}$, and also the elementary divisors of the monodromy group remain unchanged if one subjects the x -space of $x' = f(x)$ to any transformation $y = y(x)$ of the type considered in §88.

In order to prove this, notice first that the transformed Jacobi system (16), §88 possesses, by (18), §88, a fundamental matrix of the form $Y(t) = J(t)X(t)$, where $X (= R)$ is a fundamental matrix of the original Jacobi system (8), §85, and J denotes the Jacobian matrix $y_x = y_x(x)$ of the transformation $y = y(x)$ along the given periodic solution $x = x(t)$ of $x' = f(x)$; so that $J(t)$ is non-singular and has the period τ . Since $Y(t) = J(t)X(t)$, it follows that $Y(t + \tau) = J(t)X(t + \tau)$. This may be written by (5), §140, as $Y(t + \tau)$

* In fact, (2_2) then gives $D_k(t) = (tA)^k/k!$; so that (11_2) is identical with the definition of e^B in §57, if $B = tA$.

† But $A(t)$ may be periodic also when $x(t)$ is not.

$= J(t)X(t)\Gamma_X$, and so as $Y(t + \tau) = Y(t)\Gamma_X$. Comparing this with the definition (5), §140 of a monodromy matrix, one sees that $\Gamma_Y = \Gamma_X$. This implies the theorem which was to be proved.

§148. If the given periodic solution $x(t)$ of $x' = f(x)$ is not an equilibrium solution, i.e., if $x(t) \neq \text{const.}$, then at least one of the multipliers of the corresponding Jacobi system $\xi' = A(t)\xi$ is $s = 1$. This statement is, by §143, equivalent to the one that at least one characteristic exponent $\lambda = 0$. Hence, the first of the criteria of §144 bis shows that the statement is equivalent to the existence of a $\xi = \xi(t)$ which satisfies $\xi' = A(t)\xi$, has the period τ , and is of the form $\xi = \phi(t)$, where $\phi(t + \tau) = \phi(t) \neq 0$. But $x(t + \tau) = x(t) \neq \text{const.}$, by assumption; so that, by the end of §87, one can choose $\phi(t) = x'(t)$.

§149. Suppose that the given period solution $x(t) = x(t + \tau) \neq \text{const.}$ of $x' = f(x)$ may be embedded into a family of periodic solutions $x = x(t/\tau(\epsilon), \epsilon)$ of $x' = f(x)$, where $x(u, \epsilon)$ is a function of two variables which has continuous partial derivatives of the first order; and that the period $\tau = \tau(\epsilon)$, considered as a function of the integration constant ϵ (which vanishes for the embedded solution $x(t)$), has a continuous derivative $\tau_\epsilon(\epsilon)$ which does not vanish at $\epsilon = 0$. Then application of the rule (13), §87 to the family $x(t; \epsilon) = x(t/\tau(\epsilon), \epsilon)$ shows that the Jacobi system $\xi' = A(t)\xi$ belonging to $x = x(t)$ admits the solution

$$(12) \quad \begin{aligned} \xi(t) &= \psi(t) + t\phi(t), \quad \text{where} \\ \psi(t) &= x_\epsilon(t, 0), \quad \phi(t) = ax'(t), \quad a = -\tau_\epsilon(0)/\tau^2(0). \end{aligned}$$

Since $\psi(t)$ and $\phi(t)$ clearly have the period $\tau = \tau(0)$, and since $\phi(t) \neq 0$ in view of the assumptions $x(t) \neq \text{const.}$ and $\tau_\epsilon(0) \neq 0$, it follows from the second of the criteria of §144 bis, that the Jacobi system $\xi' = A(t)\xi$ has, besides the periodic solution $\xi = x'(t)$, found in §148, the secular solution (12) which belongs again to $\lambda = 0$. Thus, at least two characteristic exponents $\lambda = 0$, i.e., at least two multipliers $s = 1$.

§150. Without assuming anything else, suppose that the linear system (1₁) is canonical; so that the m -vector ξ is a $2n$ -vector, formed by n momenta and n coordinates. Thus, one has to do with the canonical system $\xi' + IH\xi = 0$ (cf. §91), in which the Hamiltonian function $H = H(\xi; t)$ is the quadratic form $\frac{1}{2}\xi \cdot H(t)\xi$ belonging to a given symmetric $2n$ -matrix $H = H(t)$. Accordingly,

(13₁) $I\xi' = H(t)\xi$, i.e., $\xi' = -IH(t)\xi$; (13₂) $H = H'$, $I' = -I = I^{-1}$;

so that $A(t) = -IH(t)$ in (1₁). According to §105 bis, the transformation of $\xi(0)$ into $\xi(t)$ is a canonical transformation of multiplier $\mu = 1$. Hence, comparison of (1₂), §137 with §37 shows that (14₁), §37 is satisfied by $\mu = 1$ and $\Gamma(t) = R(t)$, where $R(t)$ is defined by (2₁)–(2₂), §137. In other words, $R(t)$ is, for every fixed t , a completely canonical matrix (§60).

§151. Suppose, in particular, that (13₁) satisfies the assumption of §140 i.e., that $H(t + \tau) = H(t)$ for a fixed $\tau \neq 0$. According to (6) and the last remark of §150, the monodromy matrix Γ_R is completely canonical. Since the characteristic numbers and the elementary divisors of Γ_R are the invariants of the monodromy group (the former being the $m = 2n$ multipliers s_1, \dots, s_{2n} ; cf. §141–§142), it follows from §60 that if s is a multiplier, then, whether s is real or complex, s^{-1} is a multiplier and belongs, if $s \neq \pm 1$, to elementary divisors of the same degree as s (and has, in particular, the same multiplicity). In view of §143, one can express this by saying that if λ is a characteristic exponent, then so is $-\lambda$; and that if λ is neither a multiple of $2\pi i/\tau$ nor a multiple of $\pi i/\tau$, then $-\lambda$ has the same multiplicity, and belongs to secular terms of the same order, as λ . In addition, the multiplicity of $s = -1$ (i.e., of $\lambda = \pi i/\tau$), hence also the multiplicity of $s = 1$ (i.e., of $\lambda = 2\pi i/\tau$) is an even number. In fact, the product of all $2n$ multipliers s is the determinant of the completely canonical matrix Γ_R and so, by §32, equal to $+1$.

Besides the reciprocal pairing (s, s^{-1}) of the $2n$ multipliers, one has, for complex s (if any) and their multiplicities and elementary divisors, the conjugate complex pairing (s, \bar{s}) ; cf. §142. It follows, therefore, from §143 that if a characteristic exponent λ is neither real nor purely imaginary, then not only $-\lambda$ but also $\bar{\lambda}$, hence also $-\bar{\lambda}$, is a characteristic exponent; furthermore, the four distinct characteristic exponents $\pm \lambda, \pm \bar{\lambda}$ have the same multiplicities and belong to secular terms of the same order; cf. §144–§144 bis.

§152. Let $x = x(t)$ be a given solution of a conservative canonical system with n degrees of freedom. Then the corresponding Jacobi equations are, by §101, canonical, and may, therefore, be written in the form (13₁)–(13₂), §150. If, in addition, $x(t + \tau) = x(t)$, then $H(t + \tau) = H(t)$, the reason being the same as in §146. If there exists a Lagrangian function, the Hamiltonian and Lagrangian forms (21₁)–(21₂), §101 of the Jacobi equations lead to the same invariants

of the monodromy group. In fact, the passage from the Hamiltonian to the Lagrangian form of the equations of motion is, by §6–§8, a transformation of the form considered in §146. If the given periodic solution is not an equilibrium solution, then it is assured by §148 that at least one, hence by §151 that at least two, of the multipliers s is 1; so that at least two characteristic exponents λ vanish (mod $2\pi i/\tau$).

§153. Suppose finally that the given solution $x = x(t)$ is an equilibrium solution; so that $H(t) = \text{const.}$ in (13₁)–(13₂). Then the s are undefined, while the λ are uniquely determined as the characteristic numbers of $A = -IH$; cf. §145. However, the results of §151, when stated in terms of the λ , remain valid. In order to prove this, it is sufficient to show that $-A$ and A , or, what is the same thing, $-A$ and the transposed matrix A' , have the same elementary divisors, i.e., that $-A = T^{-1}A'T$ for a suitable T . But $A = -IH$; so that, by (13₂), one can choose $T = I$.

Since the matrix $R(t)$ is, by §150, completely canonical for every t , and since $R(t)$ is, in the present case, represented by (11₂), where $A = -IH$, it follows that e^{-tH} is a completely canonical matrix for every t , whenever $H = H'$. Actually, this fact has already been verified in §60 bis, since $-tH$ is symmetric for every t .

§153 bis. It should be mentioned that, while e^{tH} is a canonical matrix of multiplier $\mu = 1$ for every $H = H'$, not every canonical matrix of multiplier $\mu = 1$ may be represented by means of a suitable $H = H'$ in the form e^{tH} . The characterization of those matrices which are representable in the form e^{tH} is known and is connected with the results to which reference will be made in §154 bis.

§154. Let $F = \text{const.}$ be a symmetric $2n$ -matrix which may have a vanishing determinant but is not the zero matrix (0). If G is another such matrix, §23 and (19), §20 show that the quadratic forms $\frac{1}{2}\xi \cdot F\xi$, $\frac{1}{2}\xi \cdot G\xi$ are in involution if and only if $\xi \cdot GIF\xi \equiv 0$. This means that the matrix GIF is skew-symmetric, i.e., that $GIF = -FIG$. It follows, therefore, from §92 that the quadratic form $\frac{1}{2}\xi \cdot F\xi$ is an integral of the conservative linear canonical system $I\xi' = H\xi$ of §153 if and only if $HIF = FIF$.

This holds also when the quadratic form $\frac{1}{2}\xi \cdot F\xi$ is the square of a linear form $f \cdot \xi$, where $f = \text{const.} \neq 0$ is a $2n$ -vector. Actually, substitution of $\xi' = -IH\xi$ into $(f \cdot \xi)' \equiv f \cdot \xi'$ shows that $f \cdot \xi$ is an integral if and only if $HIf = 0$, i.e., $Hg = 0$, where $f = -Ig$; so that

$\det H = 0$ is necessary and sufficient for the existence of a linear integral $f \cdot \xi$, the number of independent linear integrals $f \cdot \xi$ being identical with the number (≥ 0) of linearly independent solutions g of the homogeneous equations $Hg = 0$, where $g = If$.

§154 bis. If $C = \text{const.}$ is any completely canonical matrix, the Hamiltonian function $\frac{1}{2}\xi \cdot H\xi$ of (13₁) is transformed by $\xi = C\eta$ into $\frac{1}{2}\eta \cdot K\eta$, where $K = C'HC$; cf. §37 and §60. Thus, there arises the question, what possible normal forms can be reached for a given $H = H' = \text{const.}$, if one replaces H by $C'HC$, where $C = \text{const.}$ is a completely canonical matrix which may be chosen suitably. This question, together with an analogous question as to the normal forms of the C themselves, has been answered generally only recently. The algebraic details involved are too lengthy to be presented here. As to canonical normal forms of certain H of particular type, cf. §64–§64 bis, where $H = Q$.

CHAPTER III

DYNAMICAL SYSTEMS

Hamiltonian and Lagrangian equations	§155–§166
Isoenergetic reduction	§167–§184
Single degree of freedom	§185–§193
Integrable systems	§194–§205
Systems with radial symmetry	§206–§226
Two degrees of freedom	§227–§240

Hamiltonian and Lagrangian Equations

§155. A Lagrangian function $L(q', q; t)$ is said to belong to a (non-relativistic) dynamical system if the Hessian n -matrix function $(L_{q'_i q'_k})$ of §15 does not contain q' and* is positive definite. These properties are invariant under the transformation of §10. By §9 bis, one can assume without loss of generality that $L_t \equiv 0$.

Thus, the Lagrangian functions in question are those and only those L which have the form

$$(1) \quad L(q', q) \equiv L = \frac{1}{2} \sum \sum g_{ik}(q) q'_i q'_k + \sum f_i(q) q'_i + U(q), \left(\sum = \sum_1^n \right),$$

where $g_{ik} = g_{ki}$, f_i , U are $\frac{1}{2}n(n+1) + n + 1$ given scalar functions of the position $q = (q_i)$ in the configuration space, and

$$(2_1) \quad T \equiv \frac{1}{2} \sum \sum g_{ik}(q) q'_i q'_k > 0 \quad \text{if} \quad \sum q_i'^2 \neq 0;$$

$$(2_2) \quad g_{ik} = L_{q'_i q'_k} = T_{q'_i q'_k} = g_{ki}.$$

According to (1) and §96 bis, the energy integral of $[L]_q = 0$ is, if h denotes the integration constant of energy,

$$(3) \quad \begin{aligned} \frac{1}{2} \sum \sum g_{ik}(q) q'_i q'_k - U(q) &= h = \text{const.}, \quad \text{i.e.,} \\ T(q', q) - U(q) &= h. \end{aligned}$$

This relation does not contain the coefficients $f_i(q)$ of the terms of (1) which are linear in the velocities. Accordingly, these terms correspond to forces which do no work, as illustrated by forces of the Coriolis type (cf. §231). Corresponding to (3), (2₁), one calls T the

* This additional assumption, i.e., (2₁), will not actually be used until §166; so that, for the present, it would be sufficient to assume that $\det (L_{q'_i q'_k}) \equiv \det (g_{ik}) \neq 0$.

kinetic, $-U$ the potential energy, while U itself is called the force function.

§156. Clearly, one does not change $[L]_q = 0$ by adding to (1), i.e., to $U(q)$, a constant, the only effect of this addition being a shift of the zero level of the energy constant (3). Furthermore, one does not change $[L]_q = 0$ by adding to every $f_i(q)$ the derivative $G_{q_i}(q)$ of a scalar $G = G(q)$. In fact, one then adds to (1) the term $\sum G_{q_i}(q)q'_i \equiv (G(q))'$ which, by the end of §94, can be omitted. Correspondingly, by the identical vanishing of all $f_i(q)$ will be meant that all $f_i(q) \equiv 0$ after a suitable choice of $G(q)$, i.e., that (f_i) is a gradient.

In this particular case, i.e., if (1) reduces to $L = T(q', q) + U(q)$, the dynamical system $[L]_q = 0$ is called of the reversible, and otherwise of the irreversible, type. The reason for this terminology is that $q = q(-t)$ is, for every solution $q = q(t)$ of $[L]_q = 0$, again a solution if and only if (1) reduces to $T + U$. This will be seen in §163.

On the other hand, $q = q(t - t^0)$ is, for every $t^0 = \text{const.}$, and for every solution $q = q(t)$ of $[L]_q = 0$, always a solution of $[L]_q = 0$ and represents the same path in the configuration space (and has, in particular, the same energy constant h) as $q = q(t)$. In fact, (1) being of the conservative type, $[L]_q = 0$ does not contain t explicitly.

§157. Since $(L_{q'_i q'_k}) \equiv (g_{ik}(q)) = (g_{ki})$ is, by (2₁)-(2₂), a positive definite matrix for every q , the assumption $\det(L_{q'_i q'_k}) \neq 0$ of §15 is satisfied. The reciprocal matrix, $(g_{ik})^{-1}$, which will be denoted by $(g^{ik}) \equiv (g^{ik}(q)) = (g^{ki})$, is again positive definite. Furthermore, from (1), §155 and (1₁)-(1₂), §15,

$$(4) \quad L_{q'_i} \equiv p_i = \sum g_{ik}q'_k + f_i, \quad \text{i.e.,} \quad H_{p_i} \equiv q'_i = \sum (p_k - f_k)g^{ik},$$

since $(g^{ik}) = (g_{ik})^{-1}$. Hence, on placing

$$(5) \quad f^i(q) \equiv f^i = \sum g^{ik}f_k, \quad \text{i.e.,} \quad f_i(q) \equiv f_i = \sum g_{ik}f^k,$$

and

$$(6) \quad \begin{aligned} V(q) &\equiv V = U - \frac{1}{2} \sum \sum g^{ik}f_i f_k, \quad \text{i.e.,} \\ U(q) &\equiv U = V + \frac{1}{2} \sum \sum g_{ik}f^i f^k, \end{aligned}$$

one sees from (2₁), §15 that (1), §155 belongs to

$$(7) \quad H(p, q) \equiv H = \frac{1}{2} \sum \sum g^{ik}(q)p_i p_k - \sum f^i(q)p_i - V(q).$$

Since $(H_{p_i p_k}) = (g^{ik})$, it follows that a conservative dynamical sys-

tem can be characterized not only by means of a Lagrangian function L which is a quadratic polynomial (1) in the velocities q'_i , with coefficient functions g_{ik} , f_i , U which depend only on $q = (q_i)$ and determine a positive definite quadratic part (2), but also by means of a Hamiltonian function H which is a quadratic polynomial (7) in the momenta p_i , with coefficients g^{ik} , f^i , V which depend only on $q = (q_i)$ and are such that

$$(8_1) \quad \frac{1}{2} \sum \sum g^{ik}(q) p_i p_k > 0 \quad \text{if} \quad \sum p_i^2 \neq 0;$$

$$(8_2) \quad T = \frac{1}{2} \sum \sum (p_i - f_i)(p_k - f_k) g^{ik}.$$

In fact, (8₁) and (8₂) are, by (4), equivalent to (2₁).

Finally, (4), (5), (6), (8₂) imply that (3), (7) can be written as

$$(9_1) \quad H(p, q) = h;$$

$$(9_2) \quad H(p, q) = T - U(q), \quad \text{where} \quad T = T(p, q), \quad \text{by} \quad (8_2).$$

§158. It is clear from (5) that the reversible case $(f_i) \equiv (0)$ of §156 can be characterized also by $(f^i) \equiv (0)$, and so, in view of (4), by

$$(10) \quad L_{q'_i} \equiv p_i = \sum g_{ik} q'_k, \quad \text{i.e.,} \quad H_{p_i} \equiv q'_i = \sum g^{ik} p_k,$$

or, on using (3), (7) and (8₂), also by

$$(11_1) \quad L(q', q) = T + U; \quad (11_2) \quad H(p, q) = T - U;$$

$$(11_3) \quad \sum \sum g_{ik} q'_i q'_k = 2T = \sum \sum g^{ik} p_i p_k.$$

Correspondingly, it is clear from (6) that $(f^i) \equiv 0$ is equivalent to $U = V$. According to (11₂), the Hamiltonian equations $q'_i = H_{p_i}$, $p'_i = -H_{q_i}$ reduce to

$$(12) \quad \begin{aligned} q'_i &= T_{p_i}, & p'_i &= U_{q_i} - T_{q_i}, \quad \text{where} \\ T &= \frac{1}{2} \sum \sum g^{ik}(q) p_i p_k, & U &= U(q). \end{aligned}$$

It is clear from (12), where $\sum p_i T_{p_i} \equiv 2T$, and from (10), that

$$(13_1) \quad \left(\sum p_i q_i \right)' = - \sum q_i (T_{q_i} - U_{q_i}) + 2T;$$

$$(13_2) \quad \sum p_i q_i = \sum \sum g_{ik} q_i q'_k.$$

§159. If, in particular, all $g_{ik} = g_{ik}(q_1, \dots, q_n)$ are homogeneous of some fixed degree α , then, since $(g_{ik})^{-1} = (g^{ik})$, the Hamiltonian kinetic energy $T = \frac{1}{2} \sum \sum g^{ik} p_i p_k$ is homogeneous of degree $-\alpha$ in the coordinates q_i , i.e., $\sum q_i T_{q_i} = -\alpha T$; and so (13₁)–(13₂), (9₁)–(9₂) show that*

* The identity (14) plays a rôle in statistical mechanics ("virial theorem").

$$(14) \quad \left(\sum \sum g_{ik} q_i q_k' \right)' = (\alpha + 2)(U + h) + \sum q_i U_{q_i}$$

is an identity in t along any solution $q = q(t)$ of energy h .

In the important particular case $\alpha = 0$, the expression on the right of (14) can be written as

$$(15_1) \quad 2(U + h) + \sum q_i U_{q_i}; \quad (15_2) \quad (\beta + 2)U + 2h;$$

$$(15_3) \quad 2(U^* + h) + \sum q_i U_{q_i}^*,$$

according as $U = U(q_1, \dots, q_n)$ is arbitrary, homogeneous of some degree β (e.g., $U \equiv 0$) or such that there exists a $U^* = U^*(q_1, \dots, q_n)$ for which $U - U^*$ is homogeneous of degree $\beta = -2$.

If α is arbitrary and U homogeneous of degree $\beta = -\alpha - 2$ (e.g., $U \equiv 0$), then (14) shows that $[L]_q = 0$ has, besides the energy integral (3), the integral†

$$(16) \quad \sum \sum g_{ik} q_i q_k' + \beta t \left(\frac{1}{2} \sum \sum g_{ik} q_i' q_k' - U \right) = \text{const.} \\ (\beta = -\alpha - 2).$$

§160. Suppose that $U(q_1, \dots, q_n)$ is homogeneous of some degree β ; or, what is, if $\beta = 0$, a more general assumption, that all $U_{q_i}(q)$ are homogeneous of some fixed degree γ ($= \beta - 1$). Suppose further that all $g_{ik}(q)$ are independent of q ; so that (12) reduces to $q_i' = \sum g^{ik} p_k$, $p_i' = U_{q_i}$, i.e., to $q_i'' = K_i(q)$, where $K_i = \sum g^{ik} U_{q_k}$. Since every $K_i = K_i(q_1, \dots, q_n)$ is homogeneous of a fixed degree γ , it is natural to seek pairs of *fixed* scalar functions $u = u(t)$, $v = v(t)$ of the time which have the property that $q_i = v(t)q_i(u(t))$ is, for every solution $q_i = q_i(t)$ of the equations of motion $q_i'' = K_i(q)$, again a solution ("dynamical similarity"). It will be assumed that $u = u(t)$, $v = v(t)$ have continuous second derivatives $u''(t)$, $v''(t)$, and that $v(t) > 0$, $u'(t) > 0$. In particular, one can introduce $u = u(t)$ instead of t as an independent variable; so that $t = t(u)$.

Since the K_i are homogeneous of degree γ ($= \beta - 1$), it is easily found by direct substitution that if $q_i = q_i(t)$ is a fixed solution of the system $q_i'' = K_i(q)$, where $q_i'' = d^2 q_i / dt^2$, then $q_i = v(t)q_i(u(t))$ is again a solution if and only if

† If every U_{q_i} is homogeneous of degree $\gamma = -1$ (for which it is sufficient but not necessary that U is homogeneous of degree $\beta = 0$), then $[L]_q = 0$ has the integral $\sum q_i U_{q_i} = \text{const.}$ (unless all $U_{q_i} \equiv 0$). This holds also when the g_{ik} are not homogeneous of some degree α , and also when the system is irreversible. In fact, $(\sum q_i U_{q_i})' \equiv 0$, since $U_{q_i} + \sum q_k U_{q_i q_k} \equiv 0$ in view of $F \equiv -\sum q_k F_{q_k}$, where $F = U_{q_i}$.

$$v^\gamma \frac{d^2 q_i}{du^2} \left(\frac{dt}{du} \right)^3 = \frac{d^2(vq_i)}{du^2} \frac{dt}{du} - \frac{d(vq_i)}{du} \frac{d^2 t}{du^2}$$

is, in virtue of $t = t(u)$, an identity in u . It follows, therefore, by comparison of the coefficients of $d^j q_i / du^j$, where $j = 0, 1, 2$, that the two functions u, v of t will have the desired property with reference to every solution $q_i = q_i(t)$ of $q_i'' = K_i$ if the two functions t, v of u satisfy the three conditions

$$(17_1) \quad v/\dot{t}^2 = v^\gamma; \quad (17_2) \quad 2\dot{v}\dot{t} - v\ddot{t} = 0; \quad (17_3) \quad \ddot{v}\dot{t} - \dot{v}\ddot{t} = 0,$$

where the dots denote differentiations with respect to u . Now, (17₃) means that \dot{v}/\dot{t} is a constant, say c . On differentiating (17₁) with respect to u , and substituting the resulting representation of \ddot{t} into (17₂), one sees from $\beta = \gamma + 1$ that the three conditions (17_k) for the two functions $t(u), v(u)$ are equivalent to

$$(18_1) \quad \dot{t}^2 = v^{2-\beta}; \quad (18_2) \quad 4\dot{t}^2\dot{v} = (2 - \beta)v^{2-\beta}\dot{v}; \quad (18_3) \quad \dot{v} = c\dot{t}.$$

Choose the integration constant $c = 0$. Then (18₃) means that the (positive) function v is a constant, say $\lambda (> 0)$; while (18₂) reduces to $0 = 0$, and (18₁) to $dt/du = \lambda^{1-\frac{1}{2}\beta}$. Consequently, all conditions are satisfied by $v = \lambda = \text{const.} > 0$ and $u = \lambda^{\frac{1}{2}\beta-1}t$; so that $q_i = \lambda q_i(\lambda^{\frac{1}{2}\beta-1}t)$ is, for every solution $q_i = q_i(t)$ of $[L]_q = 0$ and for every constant $\lambda > 0$, again a solution.

If $\beta \neq 0$, so that $U(q_1, \dots, q_n)$ is homogeneous of degree β , it is clear from (3) that the energy constant of the solution $\lambda q_i(\lambda^{\frac{1}{2}\beta-1}t)$ is λ^β times the energy constant h of $q_i(t)$.

§160 bis. If $[L]_q = 0$ has a family of periodic solutions which satisfies certain conditions of differentiability, then the period of a solution within the family is, by §100, a function $\tau = \tau(h)$ of the energy constant h alone. If the dynamical system is of the type considered in §160 and if $\beta \neq 0$, this function $\tau = \tau(h)$ can be determined explicitly.

In fact, if one extends the periodic family by introduction of the additional parameter λ , the end of §160 shows that the period and the energy constant become $\lambda^{1-\frac{1}{2}\beta}\tau(h)$ and $\lambda^\beta h$, respectively. Hence, the product $\tau(h)\lambda^{1-\frac{1}{2}\beta}$ must be a function of the product $\lambda^\beta h$, where $\lambda > 0$ is arbitrary. This means that $\tau(h)$ is, within the family, proportional to the $(\beta^{-1} - \frac{1}{2})$ -th power of $|h|$.

§161. Since the discussion of (18₁)–(18₃) in §160 was based on the

assumption $c = 0$, it remains to be seen how far are the results of §160 complete.

First, if $\beta \neq -2$, then c must be chosen to be 0. In fact, since $v > 0$ and $\dot{t} > 0$ by assumption, (18₁) implies that (18₂) cannot be satisfied for $\beta \neq -2$ unless $\dot{v} \equiv 0$, which means that $c = 0$; cf. (18₃), where $\dot{t} > 0$.

Let, however, $\beta = -2$. Then (18₂) is, in virtue of (18₁), an identity also when $\dot{v} \neq 0$; so that one can choose the constant c of (18₃) arbitrarily. Thus, the three conditions (18_k) reduce to $u'^2 = v^{\beta-2}$, $v'\dot{t} = ct$, or, since $\beta = -2$, $\dot{t} > 0$, to $u' = v^{-2}$, $v' = c$, where $u = u(t)$, $v = v(t)$. In other words, all conditions are satisfied by $u(t) = \int v(t)^{-2} dt$, $v(t) = ct + b$, where b , $c (\neq 0)$ are arbitrary constants. In particular, $q_i = \pm tq_i(1/t)$ is, for every solution $q_i = q_i(t)$ of $[L]_q = 0$, again a solution.

On comparing this situation with §96 (and §9 bis), one will expect that, corresponding to the pair b, c of arbitrary constants, there exist, if $\beta = -2$, two independent integrals which do not exist for $\beta \neq -2$. These two integrals of $[L]_q = 0$ actually exist; one of them being (16), where $\alpha = 0$, while the other, namely

$$(16 \text{ bis}) \quad \frac{1}{2} \sum \sum g_{ik} (q_i q_k - 2t q_i q'_k + t^2 q'_i q'_k) - t^2 U = \text{Const.},$$

is an obvious consequence of (16) and (3), since the g_{ik} are constants.

§162. If $U(q_1, \dots, q_n)$ is homogeneous of some degree β and the g_{ik} are independent of q (hence, $\alpha = 0$), then

$$(19_1) \quad \frac{1}{2} J'' = (\beta + 2)U + 2h; \quad (19_2) \quad J(q) \equiv J = \sum \sum g_{ik} q_i q_k.$$

In fact, (19₁) is, in view of (15₂) and of the definition (19₂), identical with (14).

If, in particular, $\beta = -2$, then (19₁) reduces to $J'' = 4h$; so that $J(t) = 2ht^2 + \text{const. } t + \text{Const.}$ This, when compared with (19₂), shows that in the exceptional case $\beta = -2$ (§161) the only solutions $q = q(t)$ of $[L]_q = 0$ which remain bounded when $t \rightarrow \pm \infty$ are those along which (19₂) is independent of t ; and that the vanishing of the energy constant h is a necessary condition for these solutions. For instance, $J(t) = \text{Const.}$ and $h = 0$ for every periodic solution, if $\beta = -2$.

§163. Returning to the general case of §155, define, in terms of the coefficients $g_{ik}(q)$, $f_i(q)$ of (1), the functions $P_{ik} = -P_{ki}$, $\Gamma_{ijk} = \Gamma_{jik}$ of the position q in the configuration space by placing

$$(20_1) \quad P_{ik} = \frac{\partial f_i}{\partial q_k} - \frac{\partial f_k}{\partial q_i}; \quad (20_2) \quad 2\Gamma_{ijk} = \frac{\partial g_{ik}}{\partial q_j} + \frac{\partial g_{jk}}{\partial q_i} - \frac{\partial g_{ij}}{\partial q_k}.$$

Then substitution of (1), §155 into (6), §94 shows that the explicit form of the system $[L]_{q_i} = 0$ of n Lagrangian equations in case of an arbitrary conservative dynamical system with n degrees of freedom is

$$(21) \quad [L]_{q_i} \equiv \sum_k g_{ik} q_k'' + \sum_j \sum_k \Gamma_{jki} q_j' q_k' + \sum_k P_{ik} q_k' - U_{q_i} = 0,$$

a system quadratic in the velocities q_i' and linear in the accelerations q_i'' . Since $(g_{ik})^{-1} = (g^{ik})$, one can solve (21) with respect to the q_i'' :

$$(22) \quad q_i'' = - \sum_j \sum_k \Gamma_{ijk}^i(q) q_j' q_k' - \sum_k P_k^i(q) q_k' + \sum_k g^{ik}(q) U_{q_k}(q),$$

where $\Gamma_{jk}^i = \sum_l g^{il} \Gamma_{ikl} = \Gamma_{kj}^i$ and $P_k^i = \sum_l g^{il} P_{lk}$ ($\neq -P_k^i$).

It should be mentioned for later application that, on assigning to a $t = t^0$ the initial conditions $q(t^0) = q^0$, $q'(t^0) = q'^0$, one has

$$(23_1) \quad q_i'(t) = q_i''(t^0)(t - t^0) + o(|t - t^0|) \text{ as } t \rightarrow t^0 \pm 0, \text{ if } q'^0 = 0;$$

$$(23_2) \quad q_i''(t^0) = \sum g^{ik}(q^0) U_{q_k}(q^0), \text{ if } q'^0 = 0.$$

In fact, (23₁) is Taylor's formula, and (23₂) a consequence of (22).

On changing t to $-t$, one sees that the n equations (21) remain unchanged if and only if all $\sum P_{ik} q_k' = 0$. This will be the case along an arbitrary solution $q = q(t)$ if and only if all $P_{ik}(q) = 0$. And (20₁) shows that all $P_{ik}(q) \equiv 0$ if and only if (f_1, \dots, f_n) is the gradient of some $G = G(q)$. This proves the statement of §156 concerning reversible systems.

§164. The n -dimensional q -domain under consideration can be thought of as carrying the Riemannian geometry determined by the covariant metric tensor (g_{ik}) . Then (f^i) and (f_i) are, by (5), the contravariant and covariant components of the same vector; while (20₂) defines the Christoffel symbols of the g_{ik} , and (20₁) the (covariant) curl of (f_i) . Furthermore, (4) shows that the momenta p_i correspond to covariant vectors (cf. §48), so that their index i is correctly written as subscript; and that the velocities q_i' correspond to contravariant vectors, so that their index i ought to be written as a superscript. In this sense, the formulae of §155 and §157 are to the effect that the Lagrangian and Hamiltonian theories are contravariant and

covariant, respectively. It is, however, clear from §158 that the kinetic energy T is an invariant of this tensor analysis (and, correspondingly, $U = V$) if and only if the system is reversible; $(f_i) \equiv (0)$ being the condition for a dynamical system in which (p_i) and (q'_i) are the covariant and contravariant representation of the same vector (cf. §15).

§165. In order to apply §79–§98 to (22), one has to replace, as in §94, the n -dimensional configuration space $q = (q_i)$ by the $2n$ -dimensional space $(q', q) = z = (z_j)$, where $z_i = q'_i$, $z_{i+n} = q_i$. For instance, an equilibrium point $q = q^0$ of the configuration space has to be defined by the property that the solution $q = q(t)$ of (22) which is determined by the initial conditions $q(t^0) = q^0$, $q'(t^0) = 0$ is $q(t) \equiv q^0$ (cf. §83). According to (22), this will be the case if and only if all n scalar sums (23₂) vanish; so that, since $\det g^{ik} \neq 0$, the equilibrium points q^0 are characterized by the vanishing of the gradient $U_q(q^0)$. It is also seen from (23₁)–(23₂) that if a solution path $q = q(t)$ reaches, as $t = t^0$, a point $q(t^0)$ of the configuration space in such a way that the velocity vector $q'(t)$ vanishes at this t^0 , then either $q'(t) \neq 0$ for every nearby t distinct from t^0 or $q'(t) \equiv 0$ according as $q(t^0)$ is not or is an equilibrium point, i.e., according as the gradient $U_q(q(t^0))$ does not or does vanish.

§166. A solution path in the (q', q) -space of §165 has, by §83, a definite tangent (and no cusp) unless the path is a single point in the (q', q) -space, i.e., an equilibrium solution. However, passage from the $2n$ -dimensional (q', q) -space to the n -dimensional q -space involves a projection, and so one cannot be sure of the existence of continuous tangents in the configuration space. Actually, it will be shown in §170 that a solution path $q = q(t)$ which is not represented by a single point of the configuration space has at any given $t = t^0$ a cusp or a definite continuous tangent according as the velocity vector $q'(t)$ does or does not vanish at $t = t^0$.

Isoenergetic Reduction

§167. With reference to an arbitrary Lagrangian function (1), §155, let \mathbf{P}_h , \mathbf{N}_h and \mathbf{Z}_h denote the sets of those points q of the n -dimensional configuration domain at which the sum of the force function $U(q)$ and of an arbitrarily fixed number h is positive, negative or zero, respectively, where it is understood that one or two, but not all three, of the sets \mathbf{P}_h , \mathbf{N}_h , \mathbf{Z}_h may contain no point q for a given h . It is clear from (3), §155, that

(i) if $q(t)$ is any solution path of energy h , then $q = q(t)$ is for every t a point of $\mathbf{P}_h + \mathbf{Z}_h$.

In fact, (2₁), §155 shows that $q(t)$ cannot be a point of \mathbf{N}_h for any t . It is also clear from (3), §155, that

(ii) if $q = q(t)$ is any solution path of energy h , the velocity vector $q'(t)$ vanishes at a given t^0 if and only if $q(t)$ is, for this t^0 , a point of \mathbf{Z}_h .

For this reason, the set \mathbf{Z}_h of points in the configuration space is called the set of zero velocity belonging to the energy level h . If $h_1 \neq h_2$, then \mathbf{Z}_{h_1} and \mathbf{Z}_{h_2} have no point in common, since

(iii) every given point, $q = q^*$, of the configuration domain is contained in exactly one \mathbf{Z}_h , namely in the one which belongs to $h = U(q^*)$. This implies, by the end of §165, that

(iv) a point q^0 represents an equilibrium solution $q(t) \equiv q^0$ of energy h if and only if $U_q(q^0) = 0$ and $U(q^0) = -h$; so that

(v) if $q = q(t)$ is a solution path of energy h and if there exists a $t = t^0$ such that the point $q(t^0)$ is on \mathbf{Z}_h and $U_q(q)$ vanishes at $q = q(t^0)$, then $q(t)$ is the equilibrium solution $q(t) \equiv q(t^0)$. Thus, (iv) shows that

(vi) if a solution path $q = q(t)$ of energy h is not an equilibrium solution, then either the point $q(t)$ is for no t on \mathbf{Z}_h , or if $q(t)$ reaches \mathbf{Z}_h when t tends to some t^0 , then the gradient $U_q(q) \neq 0$ at this point $q = q(t^0)$ of \mathbf{Z}_h .

It should be emphasized that this is true only if by "reaching \mathbf{Z}_h " is meant that $q(t)$ becomes a point of \mathbf{Z}_h for a finite $t = t^0$. In fact, it will be seen in §186 that the point $q(t)$ of a solution $q = q(t)$ which is not a equilibrium solution and is of energy h may tend to a point of \mathbf{Z}_h when $t \rightarrow \infty$. All that follows from the last remark of §165 is that

(vii) if a solution path $q = q(t)$ of energy h is such that $q(t_n)$ is a point of \mathbf{Z}_h for infinitely many distinct dates t_1, t_2, \dots , where either $t_1 < t_2 < \dots$ or $t_1 > t_2 > \dots$, then $q(t)$ is an equilibrium solution, unless $|t_n| \rightarrow \infty$ as $n \rightarrow \infty$.

§168. If q^* is any given point of the n -dimensional q -domain, and ϵ a sufficiently small positive number, let $\mathbf{Z}^\epsilon(q^*)$ denote the set of those points q at which

(1) $\mathbf{Z}^\epsilon(q^*)$: $|q - q^*| < \epsilon$ and $-U(q) = h$, where $h = -U(q^*)$, $|q - q^*|^2$ denoting $\sum (q_i - q_i^*)^2$; so that, by the definition of a \mathbf{Z}_h (§167), the set $\mathbf{Z}^\epsilon(q^*)$ is a portion (the one contained in the ϵ -neigh-

borhood about q^*) of that Z_h which contains the point q^* (cf. (iii), §167). It will always be understood that $\epsilon > 0$ is chosen sufficiently small. Two cases must be distinguished, according as the arbitrary point q^* of the configuration domain (I): is or (II): is not a zero of the gradient of the force function.

(I). Suppose first that q^* is an equilibrium point. This means by §165, that $U_q(q^*) = 0$, i.e., that the Taylor formula for $U(q) - U(q^*)$ does not contain linear terms. Hence, it is seen from the definition (1) of $Z^\epsilon(q^*)$ that the structure (dimensionality, etc.) of the set $Z^\epsilon(q^*)$ depends on the terms of higher order, the “generic” case being that $Z^\epsilon(q^*)$ consists of a finite number of $(n - 1)$ -dimensional domains which cut each other along $(n - 2)$ -dimensional subdomains of the “hypersurface” $Z^\epsilon(q^*)$. It is also seen from (1) that q^* is or is not the only point of $Z^\epsilon(q^*)$ according as $U(q)$ does or does not have at q^* an isolated extremum.

(II). On the other hand, the structure of $Z^\epsilon(q^*)$ is uniquely determined in case q^* is not an equilibrium solution. In fact, this case is, according to §165, characterized by $U_q(q^*) \neq 0$. Hence, the local existence theorem of implicit functions is applicable to (1) and shows that $Z^\epsilon(q^*)$ consists of an $(n - 1)$ -dimensional domain through q^* , does not cut itself, and has at each of its points a definite and continuous normal direction. It is also seen from (1) that, the gradient $U_q(q)$ being distinct from 0 for $q = q^*$ (hence also for $|q - q^*| < \epsilon$), the hypersurface $Z_h: U(q) = -h$ through q^* separates the ϵ -neighborhood of q^* into two n -dimensional q -domains on one of which $U(q) + h$ is positive, while on the other negative.

§169. On comparing the last remark of §168 with (i), (vi)–(vii), §167, one sees, by placing $q^* = q(t^0)$, that if a solution path $q = q(t)$ of energy h has for some $t = t^0$ a vanishing velocity vector q' , then only two cases are possible: Either

(I) the configuration path is a single point $q(t) \equiv q(t^0)$, a case characterized by the vanishing of $U_q(q)$ at the point $q = q(t^0)$ of Z_h ; or else

(II) the solution is not an equilibrium solution, i.e., $U_q(q(t^0)) \neq 0$, in which case the configuration path $q = q(t)$ will lie for $t > t^0$ on the same side of the hypersurface Z_h as it lay for $t < t^0$, it being understood that the point $q(t)$ lies on Z_h only for $t = t^0$, and that $|t - t^0|$ is supposed to be sufficiently small.

Accordingly, a solution path of energy h can never go *through* a point of Z_h , since the path either consists of a single point of Z_h or is

reflected by the hypersurface \mathbf{Z}_h (provided that it reaches \mathbf{Z}_h at all).

§170. A "reflection," just mentioned, as well as the "incidence," must take place along the transversal to the hypersurface \mathbf{Z}_h , the transversality being referred to the Riemannian metric (g_{ik}) of the configuration space (cf. §164). In other words, if there exists for a given solution path $q = q(t)$ of energy h a $t = t^0$ such that the point $q^0 = q(t^0)$ is on \mathbf{Z}_h but $U_q(q^0) \neq 0$, i.e., such that $q'(t^0) = 0$ but $q'(t) \neq 0$ for small $t - t^0 \geq 0$, then the velocity vector $q'(t)$ vanishes, as $t \rightarrow t^0 + 0$ or as $t \rightarrow t^0 - 0$, in such a way that the tangent vector, $q'(t)/|q'(t)|$, of the configuration path acquires a direction of Riemannian perpendicularity to \mathbf{Z}_h at the point q^0 of \mathbf{Z}_h .

Since the normal vector of the hypersurface \mathbf{Z}_h : $-U(q) = h$ at the point $q^0 = q(t^0)$ of \mathbf{Z}_h is $\pm U_q(q^0)/|U_q(q^0)|$, all that one has to verify is the relation

$$\frac{\left| \sum q'_i(t) U_{q_i}(q^0) \right|}{\left\{ \sum \sum g_{ik}(q^0) q'_i(t) q'_k(t) \right\}^{\frac{1}{2}} \cdot \left\{ \sum \sum g^{ik}(q^0) U_{q_i}(q^0) U_{q_k}(q^0) \right\}^{\frac{1}{2}}} \rightarrow 1;$$

$$t \rightarrow t^0 \pm 0.$$

But this relation is obvious from (23₁)–(23₂), since $(g^{ik}) = (g_{ik})^{-1}$.

The statement of §166 concerning the necessity of a cusp in case $q'(t^0) = 0 \neq q'(t)$ is, of course, a corollary; while the converse is obvious.

§171. For a fixed value of the constant h , define a Lagrangian function M by placing

$$(2) \quad \begin{aligned} M(q', q; h) &= 2T^{\frac{1}{2}}(U + h)^{\frac{1}{2}} + \sum f_i q'_i \\ &\equiv \left(\sum \sum g_{ik}(q) q'_i q'_k \right)^{\frac{1}{2}} (2U(q) + 2h)^{\frac{1}{2}} + \sum f_i(q) q'_i, \end{aligned}$$

where, as in §155, the function g_{ik}, f_i, U of $q = (q_i)$ are the coefficients of the given Lagrangian function L :

$$(3_1) \quad L(q', q) = T + \sum f_i(q) q'_i + U(q);$$

$$(3_2) \quad T = \frac{1}{2} \sum \sum g_{ik}(q) q'_i q'_k > 0, \quad \text{if } q' \neq 0;$$

so that $[L]_q = 0$ has the energy integral $T - U = h$, and so

$$(4) \quad T^{\frac{1}{2}} = (U + h)^{\frac{1}{2}} > 0, \quad \text{if } q' \neq 0, \quad \text{i.e., if } T = U + h \neq 0.$$

The meaning of $T^{\frac{1}{2}} = (U + h)^{\frac{1}{2}} (\geq 0)$ is, of course, that $T - U$ becomes a constant h along any given solution path $q = q(t)$ of $[L]_q = 0$; cf. §82.

Now consider not only *solution* paths of energy h but *any* path $q = q(t)$ which is such that, on the t -interval under consideration, the n -vector function $q(t)$ is of class $C^{(2)}$, has a non-vanishing derivative $q'(t)$, and makes (4) an identity in t for some constant h . Thus, one allows paths $q = q(t)$ which satisfy merely the energy integral $T - U = h$ of $[L]_q = 0$ for some fixed $h = \text{const.}$, without necessarily satisfying the equations of motion, $[L]_q = 0$. It will be shown that, along any such path $q = q(t)$ in the configuration space, one has, as an identity in t ,

$$(5) \quad [L]_q = [M]_q \quad \text{in virtue of} \quad (4); \quad (q' \neq 0).$$

Since $\sum f_i q'_i$ is a common additive term in the Lagrangian functions (3₁), (2), the statement (5) is equivalent to that which one obtains by writing $T + U$, $2T^{\frac{1}{2}}(2U + 2h)^{\frac{1}{2}}$ for L , M , respectively. Hence, it is seen from the definition, $[K]_q = K_{q'} - K_q$, of a Lagrangian derivative that (5) will be shown if one proves that, in virtue of (4),

$$(6) \quad \begin{aligned} T_{q'} &= \{ (2T)^{\frac{1}{2}}(2U + 2h)^{\frac{1}{2}} \}_{q'}; \\ T_q + U_q &= \{ (2T)^{\frac{1}{2}}(2U + 2h)^{\frac{1}{2}} \}_q \end{aligned}$$

(where $U_{q'} \equiv 0$, since U is a function of $q = (q_i)$ alone). Now, (3₂) shows that for the function $\{ \} = \{ (2T)^{\frac{1}{2}}(2U + 2h)^{\frac{1}{2}} \}$ of q' and q one has

$$\begin{aligned} \{ \}_{q'} &= (2T)^{-\frac{1}{2}}(2U + 2h)^{\frac{1}{2}} T_{q'}, \\ \{ \}_q &= (2T)^{-\frac{1}{2}}(2U + 2h)^{\frac{1}{2}} (T_q + U_q). \end{aligned}$$

And these relations become in virtue of (4) identical with (6), since $(2T)^{-\frac{1}{2}}(2U + 2h)^{\frac{1}{2}} \equiv 1$, $1' \equiv 0$. This proves (5).

§172. If $q = q(t)$ is any (not necessarily solution) path of class $C^{(2)}$ for which $q'(t) \neq 0$ and for which $T - U = h$ is, for some $h = \text{const.}$, an identity in t , then, on integrating the Lagrangian functions (3), (2) along the path, one has

$$(7_1) \quad S = -ht + W;$$

$$(7_2) \quad S = \int_0^t L(q', q) d\tilde{t}; \quad (7_3) \quad W = \int_0^t M(q', q; h) d\tilde{t}.$$

In fact, comparison of the definitions (7₂), (7₃) with (3₂), (2) shows that the statement (7₁) is equivalent to

$$\int_0^t (T + U) d\tilde{t} = -ht + 2 \int_0^t T^{\frac{1}{2}}(U + h)^{\frac{1}{2}} d\tilde{t},$$

a relation which clearly becomes an identity in virtue of the assumption (4). This proves (7₁) and shows also that

$$(8_1) \quad W = 2 \int_0^t T d\tilde{t} + \int_0^t \sum f_i q_i d\tilde{t}; \quad (8_2) \quad W = \int_0^t \sum p_i q'_i d\tilde{t}, \quad (p_i = L_{q'_i}),$$

(8₂) being implied by (8₁); compare (3₂), §171 with (4), §157.

On applying to both integrals (7₂), (7₃) the $\bar{\delta}$ -process mentioned at the end of §95, one sees from (7₁) that $\bar{\delta}S = \bar{\delta}W$, since $\bar{\delta}h = 0$ in view of the assumption that $T - U = h$ is a preassigned constant. This implies a new proof of (5), since it is clear from (7₂), (7₃) that $\bar{\delta}S = \bar{\delta}W$ is equivalent to (5).

The integral (7₂) is called the action, and (7₃) the isoenergetic action, belonging to the given path $q = q(t)$. It is understood that (7₂), but not (7₃), may be considered also when the path $q = q(t)$ does not satisfy $T - U = h$.

§173. The identity $\bar{\delta}S = \bar{\delta}W$, i.e., the relation (5), implies what is often referred to as the principle of Maupertuis.* What is meant is the fact, obvious from §171, that those solutions $q = q(t)$ of the Lagrangian equations $[L]_{q_i} = 0$ belonging to $L(q', q)$ which have the energy h are identical with those solutions $q = q(t)$ of the Lagrangian equations $[M]_{q_i} = 0$ belonging to $M(q', q; h)$ which satisfy the condition $T - U = h$; a condition which, in §176, will turn out to be an invariant relation of $[M]_q = 0$ (cf. §80).

Notice that this rule is applicable only in case $q'(t) \neq 0$ (cf. §171). In fact, if $q'(t) = 0$, the expression $T^{-\frac{1}{2}}$, which occurred at the end of §171, becomes meaningless; cf. (3₂). According to §169, the assumption $q'(t) \neq 0$ of Maupertuis's principle excludes, on the one hand, equilibrium solutions for every t , and, on the other hand, those t -intervals (if any) along a solution $q(t) \neq \text{const.}$ which contain a date $t = t^0$ at which the configuration path has a cusp. According to §168, both cases excluded can be characterized by the assumption that, for all t contained in the t -interval under consideration, the point $q = q(t)$ of the solution path of energy h is not on the set Z_h belonging to h .

* The actual content of this "principle" was not quite clear to Maupertuis. The precise formulation given in the text is due to Jacobi and to his predecessors, Euler and Lagrange.

§174. While the Lagrangian equations $[L]_{q_i} = 0$ can, by (22), §163, be solved with respect to the q_i' , the same does not hold for the Lagrangian equations $[M]_{q_i} = 0$. Furthermore, while there belongs to the Lagrangian function L a Hamiltonian function (cf. (7), §157) in the sense of §15, the same does not hold for the Lagrangian function M . These statements are to the effect that the Hessian $\det(M_{q'_i q'_k}) \equiv 0$. And this identity is clear from the fact that $M = M(q', q; h)$ is homogeneous of degree 1 in the n velocity components q'_i , or, more precisely, that

$$(9) \quad \lambda M(q', q; h) = M(\lambda q', q; h) \quad \text{whenever}^* \quad \lambda > 0, \quad (q' \neq 0).$$

§175. The homogeneity expressed by (9) implies that, while the Lagrangian equations $[L]_q = 0$ are, by §95, invariant under coordinate transformations, the Lagrangian equations $[M]_q = 0$ are invariant not only under coordinate transformations but under time transformations as well. In fact, if $\bar{t} = \bar{t}(t)$ is any function which has a positive continuous derivative $\bar{t}'(t)$ on the t -interval under consideration, and if one denotes by dots differentiations with respect to the new time variable \bar{t} (so that $q' = \bar{t}'\dot{q}$), then, on placing $\bar{M} = M(\dot{q}, q; h)$, one easily verifies from (9) that $[M]_q = \bar{t}'[\bar{M}]_q$ in virtue of $t = t(\bar{t})$.

§176. The last remark of §175 agrees with the first remark of §174 and shows that, when applying the rule of §173, one has to proceed as follows:

If a solution $q = q(\bar{t})$ of $[\bar{M}]_q = 0$, where $\bar{M} = \bar{M}(\dot{q}, q; h)$ and $\dot{q} = dq/d\bar{t}$, is known in terms of some given time variable \bar{t} , then \bar{t} in itself cannot be distinguished from t . In fact, the corresponding solution $q = q(t)$ of $[M]_q = 0$, where $M = M(q', q; h)$ and $q' = dq/dt$, can always be obtained from the connection $t = t(\bar{t})$ between \bar{t} and t . And this connection can always be determined from the requirement of §173 according to which the energy condition $T - U = h$ must be satisfied, if the time variable is t . In fact, (3₂) shows that $T - U = h$, i.e., (4), can be written in the form

$$(10) \quad \dot{t} \equiv \frac{dt}{d\bar{t}} = \frac{\left\{ \sum \sum g_{ik}(q) \dot{q}_i \dot{q}_k \right\}^{\frac{1}{2}}}{\left\{ 2(U(q) + h) \right\}^{\frac{1}{2}}} > 0; \quad (\dot{q} = q'\dot{t}, q' \neq 0).$$

* If $\lambda < 0$, one has to replace λ on the left of (9) by $-\lambda$, since the square roots occurring in (2), (4) have been chosen to be positive (they could have been chosen to be negative, but they cannot be chosen sometimes positive and sometimes negative, since $T^{\frac{1}{2}}$ and M cease to be of class $C^{(2)}$ when $T' = 0$, i.e., when $q' = 0$).

Now, if a solution $q = q(\bar{t})$ of $[\bar{M}]_q = 0$ is known, the connection $\bar{t} = \bar{t}(t)$ between t and \bar{t} follows from (10) by the inversion of a quadrature. In particular, the function $t = t(\bar{t})$ is uniquely determined up to an additive constant.

§177. It should be mentioned for later application that the location of the conjugate points alone determines whether an unbroken extremal of the calculus of variations problem $\bar{\delta}W = 0$ does or does not yield a proper strong minimum of (7₃); and that the same situation holds for the problem $\bar{\delta}S = 0$ belonging to (7₂). In fact, both problems satisfy the \mathcal{E} -condition in its strictest form.

First, if $Q(r, s) = \sum \sum g_{ik} r_i s_k$, the square of $Q(r, s)$ is, by (3₂) less than the product $Q(r, r)Q(s, s)$, unless the two n -vectors (r_i) , (s_i) are such that $\mu r_i = \nu s_i$ holds for a suitable pair of scalars μ, ν which are independent of i . This means, in view of (2), that

$$M(r, q; h) - M(s, q; h) - \sum (r_i - s_i) M_{s_i}(s, q; h) > 0,$$

unless the vectors r, s are proportional. Thus the Lagrangian function M , which is of the homogeneous type (9), satisfies the \mathcal{E} -condition in its strictest form.

The corresponding condition for the inhomogeneous Lagrangian function (3₁) is that

$$L(r, q) - L(s, q) - \sum (r_i - s_i) L_{s_i}(s, q) > 0,$$

unless $r_i = s_i$ for every i . According to (3₁) and (3₂), this condition is satisfied if, on placing again $Q(r, s) = \sum \sum g_{ik} r_i s_k$, one has $Q(r, s) < \frac{1}{2}Q(r, r) + \frac{1}{2}Q(s, s)$, unless $r = s$; i.e., if $Q(r - s, r - s) > 0$, unless $r = s$. Now, the assumption (3₂) is that $Q(u, u) > 0$, unless $u = 0$; so that the proof is complete.

§178. Suppose that $L = T$, i.e., that (3₁) is of the reversible type ($f_i \equiv 0$) and also that the force function $U \equiv 0$. Then $L = H$, by (11₁)–(11₃), §158; so that the energy integral is $T = h$. The Lagrangian equations (22), §163 reduce to $q_i'' = - \sum \sum \Gamma_{jk}^i q_j' q_k'$, i.e., to the equations of the geodesics on the Riemannian manifolds defined by $ds^2 = \sum \sum g_{ik} dq_i dq_k$. Thus, $T = \frac{1}{2}s'^2$, by (3₂); so that $\frac{1}{2}s'^2 = h$. In other words, $s = (2h)^{\frac{1}{2}}t$, if the arc length s on the geodesic is measured from $t = 0$ in the direction of increasing t . Correspondingly, the Lagrangian function (2) reduces to $M = (2h)^{\frac{1}{2}}s'$. Clearly, the arbitrary time variable, \bar{t} , of §175 is the arc length if and only if $\bar{M} = \dot{s}$; so that the time t becomes the arc length s if $s' = 1$, in which case $2h = 1$ and $M = s' = (2T)^{\frac{1}{2}}$.

§179. In order to generalize the assumption of §178, suppose only that L is of the reversible type, i.e., $(f_i) \equiv (0)$. Then (2) reduces to

$$(11) \quad M = (2U(q) + 2h)^{\frac{1}{2}}(2T)^{\frac{1}{2}}, \text{ where } T = \frac{1}{2} \sum \sum g_{ik}(q) q'_i q'_k.$$

This can be written in the form

$$(12) \quad M = (2\tilde{T})^{\frac{1}{2}}, \text{ where } \tilde{T} = \frac{1}{2} \sum \sum \tilde{g}_{ik}(q; h) q'_i q'_k, \tilde{g}_{ik} = 2(U + h)g_{ik}.$$

But (12) is the function (2) which belongs to the Lagrangian function \tilde{L} of geodesic type, $\tilde{L} = \tilde{T} - \tilde{U} \equiv \tilde{T}$, in the same way as the function (11) belongs to the arbitrary Lagrangian function L of the reversible type, $L = T - U$. Since the functions M defined by (11) and (12) are identical, it follows that, barring the cases of equilibrium solutions and of solution paths with cusps (§173), the reversible dynamical system is, for a fixed value of the energy constant h , equivalent to the problem of geodesics on the Riemannian manifold

$$(13) \quad d\tilde{s}^2 = \sum \sum \tilde{g}_{ik} dq_i dq_k \equiv 2(U + h)ds^2, \text{ where } ds^2 = \sum \sum g_{ik} dq_i dq_k.$$

§180. One can interpret §176 as supplying a rule for the introduction of new time variables into a dynamical system, if only solutions of a fixed energy h are considered. To the same end, one can proceed in a more direct manner, by using the Hamiltonian form of the equations.

In fact, let $G(x)$ be any continuous non-vanishing scalar function in the $2n$ -dimensional x -domain of a conservative Hamiltonian system $Ix' = H_x(x)$. Along any given solution $x = x(t)$ of $Ix' = H_x$, consider the new time variable \bar{t} defined by

$$(14) \quad \bar{t} \equiv \bar{t}(t) = \int^t \frac{dt^*}{G(x(t^*))}, \quad (G \neq 0),$$

and denote by a dot differentiation with respect to \bar{t} ; so that $\dot{\bar{t}} = 1/\bar{t}' = G$, and so $\dot{x} = x'G$. Consider only those solutions $x = x(t)$ of $Ix' = H_x$ which have a fixed energy constant h , and define a conservative Hamiltonian function \bar{H} by placing

$$(15) \quad \bar{H}(x; h) \equiv \bar{H} = (-h + H)G, \text{ where } H = H(x), G = G(x) \neq 0;$$

so that $\bar{H}_x \equiv (-h_x + H_x)G + 0 \equiv H_x G$, since $-h + H = 0$ along the solution $x = x(t)$ under consideration, and $h = \text{const.}$

It is clear from $\dot{x} = x'G$ and $\bar{H}_x = H_x G$, where $G \neq 0$, that those solutions $x = x(t)$ of $Ix' = H_x$ which have the energy constant h are, in virtue of the time transformation (14) or its inverse

$$(16) \quad t \equiv t(\bar{t}) = \int^{\bar{t}} G(x(\bar{t}^*)) d\bar{t}^*, \quad (G \neq 0),$$

identical with those solutions $x = x(\bar{t})$ of $I\dot{x} = \bar{H}_x$ which have the energy constant $\bar{h} = 0$, i.e., which satisfy the invariant relation $\bar{H} = 0$ of $I\dot{x} = \bar{H}_x$. In fact, $H = h$ and $\bar{H} = 0$ are equivalent, since $G \neq 0$ in (15).

§181. The practical merit of the rule of §180 lies in the fact that the transition (15) from $H(x)$ to $\bar{H}(x; h)$ does not involve heavy calculations, no matter how one chooses $G(x)$.

Suppose, for instance, that, writing $I\dot{x}' = H_x(x)$ more explicitly as

$$(17) \quad \begin{aligned} p'_i &= -H_{q_i}(p_1, \dots, p_n, q_1, \dots, q_n), \\ q'_i &= H_{p_i}(p_1, \dots, p_n, q_1, \dots, q_n), \end{aligned}$$

($i = 1, \dots, n$), one has $H_{p_n}(p, q) \neq 0$ in the $2n$ -dimensional (p, q) -domain under consideration. Then one can choose $G(x) \equiv G(p, q)$ to be $1/H_{p_n}(p, q)$, in which case the time variable (14) becomes $\bar{t} = q_n + \text{const.}$, since $H_{p_n} = q'_n$, by (17). Furthermore, the assumption $H_{p_n}(p, q) \neq 0$ implies that one can solve, with respect to p_n , the equation $-h + H(p_1, \dots, p_n, q_1, \dots, q_n) = 0$ in the vicinity of every point $(p_i, q_i) \equiv (p_i(t^0), q_i(t^0))$ of the given phase path $p_i = p_i(t)$, $q_i = q_i(t)$ of energy h ; so that $p_n = -K(p_1, \dots, p_{n-1}, q_1, \dots, q_{n-1}, q_n; h)$, where K is, for fixed h , a function of $2n - 1$ variables which is locally unique and satisfies the same differentiability conditions as $H(p, q)$.

Since $q_n = \bar{t}$ up to an additive integration constant which can be omitted, comparison of §180 with the definition of K shows, after straightforward reductions, that those solutions of the conservative system $\dot{p}_i = -\bar{H}_{q_i}$, $\dot{q}_i = \bar{H}_{p_i}$ with n degrees of freedom which satisfy the invariant relation $\bar{H}(p, q; h) = 0$ of §180, are identical with those solutions of the non-conservative system with $n - 1$ degrees of freedom,

$$(18) \quad \begin{aligned} \dot{p}_j &= -K_{q_j}(p_1, \dots, p_{n-1}, q_1, \dots, q_{n-1}, \bar{t}; h), \\ \dot{q}_j &= K_{p_j}(p_1, \dots, p_{n-1}, q_1, \dots, q_{n-1}, \bar{t}; h) \end{aligned}$$

($j = 1, \dots, n - 1$) which are unrestricted by any invariant system. It follows, therefore, from §180 that those solutions $p_i = p_i(t)$, $q_i = q_i(t)$; $i = 1, \dots, n$, of (17) which have the energy h are, in virtue of $\bar{t} = q_n$, identical with the solutions $p_j = p_j(\bar{t})$, $q_j = q_j(\bar{t})$; $j = 1, \dots, n - 1$, of (18).

It is, however, understood that the operation leading from (17) to (18) is of a local nature, since, in the construction of K , use has been made of the local existence theorem of implicit functions.

Needless to say, one could have introduced $\bar{t} = q_n$ by the rule of §175 also.

Clearly, one can replace the assumption $H_{p_n}(p, q) \neq 0$ by any of the $2n$ assumptions $H_{p_i}(p, q) \neq 0$, $H_{q_i}(p, q) \neq 0$, in which case \bar{t} becomes q_i , p_i , respectively. Notice that at least one of these $2n$ assumptions is satisfied in a $2n$ -dimensional vicinity of all those solution paths $x = x(t)$ of $Ix' = H_x(x)$ which are not equilibrium solutions $x(t) = \text{const.}$

§182. The above reduction of the degree of freedom from n to $n - 1$ for any fixed value of the energy constant h can, in view of §93 or §9 bis, be interpreted as the elimination of a coordinate which does not occur explicitly in the Hamiltonian function. Hence, it has to be expected that if only the time derivative of one of the n coordinates q_i , say that of q_n , occurs in (3₁), while the coefficient functions g_{ik}, f_i, U of $L(q', q)$ are independent of q_n , then one can replace the conservative dynamical system $[L]_q = 0$ with n degrees of freedom by a conservative dynamical system $[L^*]_{q^*} = 0$ with $n - 1$ degrees of freedom, where $q = (q_i)$, $i = 1, \dots, n$ and $q^* = (q_j)$, $j = 1, \dots, n - 1$. Of course, L^* must contain an integration constant (which corresponds to the fixed value of h in §181); and, if a solution $q^* = q^*(t)$ of $[L^*]_{q^*} = 0$ is known, the determination of the ignored coordinate $q_n = q_n(t)$ may be expected to require a quadrature (which corresponds to (14), a quadrature which, in §181, degenerated into $\bar{t} = q_n + \text{const.}$). This programme can easily be carried out, as follows:

If one calls a coordinate an “ignorable” (or “cyclic”) coordinate when only its time derivative occurs in L , it is clear from §15 that a coordinate is ignorable if and only if its canonically conjugate momentum, but not the coordinate itself, occurs in H . Now, if $H(p, q; t)$ is of the form $H(p_n, p^*, q^*; t)$, where the $(n - 1)$ -vectors p^*, q^* represent the first $n - 1$ components of the n -vectors p, q , then $p' = -H_q, q' = H_p$ shows that $p_n = c$, where c is an integration constant; and that $q_n = q_n(t)$ follows from $q_n' = H_{p_n}(c, p^*(t), q^*(t); t)$ by a quadrature, if one knows a solution $p^* = p^*(t), q^* = q^*(t)$ of $p^{*'} = -H_{q^*}^*, q^{*'} = H_{p^*}^*$, where $H^* \equiv H^*(c, p^*, q^*; t)$ denotes, for a fixed value of the integration constant c , the Hamiltonian function $(H(p_n, p^*, q^*; t))^{p_n=c}$ with $n - 1$ degrees of freedom.

Finally, L^* follows from H^* by the rule of §15, if one knows that the Hessian involved does not vanish.

§183. In the case of a dynamical system of the type considered since §155, the process just described may be carried out explicitly, as follows:

In order to simplify the formulae, suppose that the problem is of the reversible type, i.e., that $(f_i) \equiv (0)$. Since q_n is supposed to be an ignorable coordinate, the Lagrangian function (1), §155 becomes

$$(19) \quad L(q_n', q^{*'}, q^*) = \frac{1}{2} \sum \sum g_{ik}(q^*) q_i' q_k' + U(q^*),$$

where the summations run from 1 to n , while $j = 1, \dots, n-1$ in $q^* = (q_j)$. It follows, therefore, from (4)–(7), §157 that the Hamiltonian function with $n-1$ degrees of freedom which in §182 was denoted, for a fixed value of $c (= p_n \equiv L_{q_n'})$, by H^* , may be written explicitly as

$$(20) \quad \begin{aligned} & H^*(c, p^*, q^*) \\ &= \frac{1}{2} \sum^* \sum^* g^{il}(q^*) p_i p_l + c \sum^* g^{in}(q^*) p_i - \{ U(q^*) - \frac{1}{2} c^2 g^{nn}(q^*) \}, \end{aligned}$$

if the mark^{*} of \sum^* means that the summation runs from 1 to $n-1$. Now, (20) is of the form (7), §157, if one replaces n by $n-1$ and puts $f^i = -c g^{in}$, $V = U - \frac{1}{2} c^2 g^{nn}$. Hence, the formulae (4)–(6), §157, which define the transition from (7), §157 to (1), §155, show that the Lagrangian function with $n-1$ degrees of freedom which belongs to the Hamiltonian function (20) is given by

$$(21) \quad \begin{aligned} & L^*(c, q^{*'}, q^*) \\ &= \frac{1}{2} \sum^* \sum^* g_{il}^*(q^*) q_i' q_l' + \sum^* f_i^*(c, q^*) q_i' + U^*(c, q^*), \end{aligned}$$

where $U^* = U - \frac{1}{2} c^2 g^{nn} + \frac{1}{2} c^2 \sum^* \sum^* g^{in} g^{ln} g_{jj}^*$, $f_i^* = -c \sum^* g^{ln} g_{jl}^*$, $(g_{jl}^*) = (g^{jl})^{-1}$. The last condition defines a positive definite $(n-1)$ -matrix function (g_{jl}^*) , since, the n -matrix $(g^{ik}) = (g_{ik})^{-1}$ being positive definite by (21), §155, the same holds for the $(n-1)$ -matrix (g^{il}) .

Notice that the conservative Lagrangian function (21) with $n-1$ degrees of freedom is, in general, of the irreversible type $(f_i^*) \not\equiv 0$, although it belongs to the Lagrangian function (19) of the reversible type $(f_i) \equiv 0$.

§184. Let, in particular, $n = 2$, and, for simplicity, $g_{12} \equiv 0$; so that (20) becomes

$$L = \frac{1}{2} \sum_{i=1}^2 g_{ii}(q_1) q_i'^2 + U(q_1).$$

Then

$$(22) \quad L^* = \frac{1}{2} q_1'^2 g_{11}(q_1) + \{ U(q_1) - \frac{1}{2} c^2 / g_{22}(q_1) \},$$

by (21). Thus L^* is of the reversible type.

This is not a coincidence. In fact, every (conservative) dynamical problem with a single degree of freedom is of the reversible type. For if $q = q_1$, then (1), §155 becomes $L = \frac{1}{2} g(q) q'^2 + f(q) q' + U(q)$, where the letters denote scalars. Hence, $f(q)$ is the derivative of a function ($= \int f(q) dq$), and so §156 shows that L may be replaced by $\frac{1}{2} g(q) q'^2 + U(q)$.

Single Degree of Freedom

§185. Suppose that $n = 1$, so that $q = q_1$ is a scalar, and let, for simplicity, the q -domain be the whole q -axis. Since the system is, by §184, necessary reversible, $L = \frac{1}{2} g q'^2 + U$, where $g = g(q) > 0$, by (2₁), §155. The energy integral (3), §155, is $\frac{1}{2} g(q) q'^2 - U(q) = h$; so that

$$(1_1) \quad q'^2 = F(q; h); \quad (1_2) \quad F = 2(U(q) + h)/g(q), \quad \text{where } g(q) > 0.$$

Thus, it is seen from §167–§170 that the points $\bar{q} = \bar{q}(h)$ of the set Z_h on the q -axis can be characterized as the roots q (if any) of the equation $F(q; h) = 0$; and that if a solution $q = q(t)$ of energy h is such that $\bar{q} = q(\bar{t})$ becomes, for some $t = \bar{t}$, a root $q = \bar{q} \equiv \bar{q}(h)$ of $F(q; h) = 0$, then the solution is the equilibrium solution $q(t) \equiv \bar{q}$ or has at $t = \bar{t}$ a cusp according as the partial derivative $F_q(q; h)$ does or does not vanish at $q = \bar{q}$, i.e., according as the root \bar{q} of $F(q; h) = 0$ is multiple or simple. The first case can, by (1₂), also be characterized by $U_q(\bar{q}) = 0$; while in the second case the cusp of the solution path $q = q(t)$ is manifest from §169, since the path, which is on the q -axis, must be reflected by the point $q = \bar{q}$. Correspondingly, $q(t)$ is steadily increasing or steadily decreasing on t -intervals not containing dates of cusps, since on such t -intervals $q'^2(t) \neq 0$, and so, for reasons of continuity, either $q'(t) > 0$ or $q'(t) < 0$.

§186. On comparing the last remark of §185 with the uniqueness of the initial value problem of ordinary differential equations, and noting that $q(t + \text{const.})$ is, for any const., a solution of the same energy h as $q(t)$, one sees from (1₁) that, if $q = q_I(t)$ and $q = q_{II}(t)$

are two solution paths which have the same energy h and for which there exist two dates t_I, t_{II} on the t -axis and a point q^* on the q -axis such that $q_I(t_I) = q^* = q_{II}(t_{II})$ but $U_q(q^*) \neq 0$, then the two solution paths are identical, in the sense that $q_I(t) = q_{II}(t + \text{const.})$ for a suitable const. (Needless to say, this is a property characteristic of the case of a configuration space of dimension number $n = 1$).

Correspondingly, $[L]_q = 0$ can be solved for any preassigned value of h as follows: Exclude, for a fixed h , those points on the q -axis at which the function (1_2) is negative [cf. (1_1)], and, barring the trivial case $F(q_0; h) = 0, F_q(q_0; h) = 0$ of an equilibrium solution $q(t) \equiv q_0$, mark on the q -axis those (necessarily open) intervals (if any), at which $F(q; h)$ is positive. Then, if $I = I(h)$ is one of these intervals, and q^* an arbitrary point of I , local inversion of the quadrature

$$(2) \quad t - t^* = \int_{q^*}^q \pm |F(\tilde{q}; h)|^{-\frac{1}{2}} d\tilde{q} \quad (t^* = \text{arb. const.})$$

supplies all those solutions $q = q(t)$ of $[L]_q = 0$ which are such that the point $q(t)$ is for some value of t a point of I . This is clear from (1_1) ; while §185 shows that $q(t)$ then is for every t a point of the closure of the open interval I . It is understood that either or both of the end points of $I = I(h)$ can be at infinity, and that $F(q; h)$ must vanish at a finite end point of I (if any).

If \bar{q} is either of the two ends of I (if any), then two cases are possible, according as $F(q; h)$ vanishes at $q = \bar{q} (\neq \pm \infty)$ only in the first or in a higher order, i.e., according as $F_q(\bar{q}; h)$ does not or does vanish. In the first, but not in the second, case the solution path $q = q(t)$ which is contained in the closure of I reaches the finite end $q = \bar{q}$ of I at a finite $t = \bar{t}$. This becomes clear by letting the variable end q of the integral (2) tend to \bar{q} and then noting that the integrand becomes infinite at \bar{q} in an integrable order ($= \frac{1}{2}$) in the first case but in a non-integrable order (≥ 1) in the second case. According to §169, one has to do with the first or the second case according as the root $q = \bar{q} = \bar{q}(h)$ of $F(q, h) = 0$ is not or is an equilibrium point. Finally, it is clear† from (2) that in the second case

† Notice, however, that a corresponding remark does not hold in case the end point $q = \bar{q}$ of I is $\bar{q} = +\infty$ or $\bar{q} = -\infty$, instead of being a finite \bar{q} . If, for instance, $L(q', q) = \frac{1}{2}(q'^2 + q^4)$ and $h = 0$, then (1_2) reduces to $F(q; h) = q^4$; so that the interval $q_I < q < q_{II}$, where $q_I = 0, q_{II} = +\infty$, is an I . On the other hand, (1_1) , i.e., $q'^2 = q^4$, has the solution $q(t) = (t^0 - t)^{-1}$ for

$q(t)$ tends, when either $t \rightarrow -\infty$ or $t \rightarrow +\infty$, to the finite limit \bar{q} in such a way that $q'(t)$ does not vanish for sufficiently large $t > 0$ or $t < 0$; so that the solution $q = q(t)$ is asymptotic to the equilibrium solution which is represented by the point \bar{q} .

§187. It follows that in order to obtain solutions $q = q(t)$ which are neither equilibrium solutions nor of the asymptotic type, one has to assume that the q -interval I belonging to $q = q(t)$ is a finite interval $q_I(h) < q < q_{II}(h)$ such that $F(q; h)$ vanishes at both end points of I exactly in the first order and is positive on I .

Placing, on these assumptions, $\alpha \equiv \alpha(h) = q_I(h)$, $\beta \equiv \beta(h) = q_{II}(h)$, one sees from §185 that α is the minimum and β the maximum of $q(t)$ for $-\infty < t < +\infty$, and that $q'(t) = 0$ at those and only those t for which either $q(t) = \alpha$ or $q(t) = \beta$. Choosing, without loss of generality, the origin of the t -axis so that $q(0) = \alpha$, and noting that, the system being reversible (§184), the function $q(-t)$ also is a solution, one has $q(t) = q(-t)$. In fact, the initial values of the coordinate and the velocity determine the solution uniquely; while these initial values assigned for $q(t)$ and $q(-t)$ are identical, since $q'(0) = 0$. Furthermore, on placing*

$$(3) \quad \tau \equiv \tau(h) = 2 \int_{\alpha}^{\beta} [F(q; h)]^{-\frac{1}{2}} dq, \quad \text{where} \quad \alpha = \alpha(h), \beta = \beta(h),$$

one sees from (2) that the amount of time needed to reach $q = \beta$ from $q = \alpha$ (or $q = \alpha$ from $q = \beta$) is $\frac{1}{2}\tau$. Since $q = q(t + \text{const.})$ is the same solution path as $q = q(t)$, it follows, again from the uniqueness of the initial value problem, that not only $q(t) = q(-t)$ holds but also $q(t + \tau) = q(t)$.

Accordingly, $q(t)$ is an even periodic function which has (3) as primitive period, and α and β as its minimum and maximum, respectively.

§188. For values of t at which neither $q(t) = \alpha$ nor $q(t) = \beta$ (i.e., for values of t distinct from $\frac{1}{2}k\tau$, where $k = 0, \pm 1, \pm 2, \dots$), one

$-\infty < t < t^0 = \text{arb. const.}$; so that $q(t)$ tends to $q_{II} = +\infty$ when $t \rightarrow t^0 - 0$ and not, as one might have expected, when $t \rightarrow -\infty$.

In order to exclude this situation, one has, by (2), to assume that, for the given value of h ,

$$\int^{\pm\infty} |F(q; h)|^{-\frac{1}{2}} dq = \pm\infty; \text{ e.g., that } 0 < F(q; h) < \text{Const. } q^2, \text{ as } q \rightarrow \pm\infty.$$

* The integral (3) has a finite positive value, since $F(q; h)$ is positive for $\alpha < q < \beta$ and vanishes at $q = \alpha$ and $q = \beta$ only in the first order.

has to choose in (2) the upper or the lower sign according as $q'(t) > 0$ or $q'(t) < 0$ (i.e., according as $k\tau < t < (k + \frac{1}{2})\tau$ or $(k + \frac{1}{2})\tau < t < (k + 1)\tau$, where $k = 0, \pm 1, \pm 2, \dots$).

Instead of carrying out the periodic inversion problem assigned by (2), one usually prefers the reduction of the problem to the trivial periodic inversion problem which belongs to the linear oscillator $d^2q/d\bar{t}^2 + q = 0$, where $\bar{t} = \bar{t}(t)$ is a new time variable, to be chosen in such a way as to uniformize the multi-valued relation between t and its single-valued periodic function $q = q(t)$.

To this end, keep the energy constant h fixed and put

$$(4) \quad G(q) = [(\beta - q)(q - \alpha)/F(q; h)]^{\frac{1}{2}}$$

for $\alpha < q < \beta$. The assumptions made at the beginning of §187 are to the effect that the limits $G(\alpha + 0)$, $G(\beta - 0)$ exist and, when denoted by $G(\alpha)$, $G(\beta)$, are such that

$$(5) \quad 0 < \text{const.} \leq G(q) \leq \text{Const.} < +\infty \quad \text{for} \quad \alpha \leq q \leq \beta.$$

Since $\alpha \leq q(t) \leq \beta$ for $-\infty < t < +\infty$, it follows that one can identify (4) with the G of §180, and so introduce, along the given solution $q = q(t)$, a new time variable (14), §180. Choosing, without loss of generality, the origin of the \bar{t} -axis so that $\bar{t}(0) = 0$, one has

$$(6) \quad t \equiv t(\bar{t}) = \int_0^{\bar{t}} G(q(\bar{t}^*)) d\bar{t}^*, \quad \text{by (16), §180.}$$

If one denotes by dots and primes differentiations with respect to \bar{t} and t , respectively, $\dot{\bar{t}}$ and $\bar{t}' = \dot{\bar{t}}^{-1}$ remain, by (6) and (5), between fixed positive bounds, and so \bar{t} runs with t from $-\infty$ to $+\infty$ in a strictly increasing way. Now, \bar{t} is a uniformizing variable of the (real) relation (2) between q and t ; in addition, the system reduces in terms of \bar{t} to a linear oscillator.

In fact, it is clear from (4) and (6) that (1₁) can be written as $\dot{q}^2 = (\beta - q)(q - \alpha)$, where $\dot{q} = dq/d\bar{t}$. That solution of this differential equation which satisfies the assigned initial condition $q(0) = \alpha$ is

$$(7) \quad q = \frac{1}{2}(\beta + \alpha) - \frac{1}{2}(\beta - \alpha) \cos \bar{t}.$$

Hence, $G(q(\bar{t}))$ is an even periodic function which, when expanded into a Fourier series

$$(8_1) \quad G(q(\bar{t})) = \sum_{n=-\infty}^{+\infty} \nu_n \cos n\bar{t}; \quad (8_2) \quad \nu_n = \frac{1}{\pi} \int_0^\pi G \cos n\bar{t} d\bar{t} = \nu_{-n}$$

and substituted into (6), shows that

$$(9_1) \quad t = \nu_0 \bar{t} + \sum_{n=1}^{\infty} \lambda_n \sin n\bar{t}; \quad (9_2) \quad \lambda_n = 2\nu_n/n.$$

Since \bar{t} runs with t from $-\infty$ to $+\infty$ in a strictly monotone manner, (7) and (9₁) form a uniformization of the relation (2) between q and t , with \bar{t} as uniformizing parameter.

Since q , when considered as a function of t , is even and has the period (3), there is also a Fourier expansion

$$(10_1) \quad q = \sum_{n=-\infty}^{\infty} \rho_n \cos (nt/\nu_0);$$

$$(10_2) \quad \rho_n = \tau^{-1} \int_0^{\tau} q \cos (nt/\nu_0) dt = \rho_{-n}; \quad (10_3) \quad \tau = 2\pi\nu_0,$$

(10₃) being implied by (7) and (9₁).

§189. Suppose that the energy constant h is varying in the vicinity of a fixed 0h which is such that, while the conditions of §187–§188 are satisfied for $h \neq {}^0h$, the two subsequent simple roots $q = \alpha \equiv \alpha(h) = \min q(t; h)$, $q = \beta \equiv \beta(h) = \max q(t; h)$ of the equation $F(q; h) = 0$, where $F > 0$ for $\alpha < q < \beta$, coincide at a double root 0q of $F({}^0q; {}^0h) = 0$ when $h \rightarrow {}^0h$. Thus, there is for $h \neq {}^0h$ a definite period $\tau = \tau(h)$ represented by (3); while $\tau({}^0h)$ does not exist, since the assumption $F_q({}^0q; {}^0h) = 0$ implies that the solution $q(t; h)$ becomes for $h = {}^0h$ the equilibrium solution represented by the point 0q . Nevertheless, $\tau(h)$ tends, as $h \rightarrow {}^0h$, to a finite positive limit:

$$(11) \quad \tau(h) \rightarrow 2\pi/\sqrt{\{-\frac{1}{2}F_{qq}({}^0q; {}^0h)\}} \quad \text{as } h \rightarrow {}^0h.$$

First, the assumption $F_{qq}({}^0q; {}^0h) \neq 0$ implies that $F_{qq}({}^0q; {}^0h) < 0$. In fact, Taylor's formula shows that the ratio of the positive functions $F(q; h)$ and $(\beta - q)(q - \alpha)$, where $\alpha < q < \beta$ and $F_q({}^0q; {}^0h) = 0$, tends to the constant $-\frac{1}{2}F_{qq}({}^0q; {}^0h)$ as $h \rightarrow {}^0h$, i.e., as $|\alpha - \beta| \rightarrow 0$. Hence, (11) is clear from (3).

If $g(q) \equiv 1$, the limit (11) becomes $2\pi/\sqrt{\{-U_{qq}({}^0q)\}}$, by (1₂).

§190. The periodicity assumption of §187–§188 will now be omitted. Suppose* that $g(q) \equiv 1$, i.e., that $L(q', q) = \frac{1}{2}q'^2 + U(q)$;

* This supposition involves, for a fixed value of the energy constant h , no loss of generality, as is seen by identifying the function $G(x)$ of §180 with the given positive function $g(q)$, and then applying the transformation of §180 to the Hamiltonian function $H = \frac{1}{2}g^{-1}p^2 - U$ which, by §158, belongs to $L = \frac{1}{2}gq'^2 + U$.

cf. §185. Then $[L]_q \equiv q'' - U_q(q)$, and so §101 shows that the Jacobi equation determining the displacements $\kappa = \kappa(t)$ of a given solution $q = q(t)$ of $[L]_q = 0$ is $\kappa'' + a(t)\kappa = 0$, where $a(t) = -U_{qq}(q(t))$. Hence, the coefficient $a(t)$ is periodic on the assumption of §187–§188; while it becomes the constant $-U_{qq}({}^0q)$ in case $q(t)$ is an equilibrium solution $q(t) \equiv {}^0q$. In the latter case, the characteristic exponents (§89) of the Jacobi equation are $\pm \sqrt{U_{qq}({}^0q)}$. And the general solution of $\kappa'' + a\kappa = 0$ is an hyperbolic or a linear function of t according as $a = -U_{qq}({}^0q)$ is negative or zero; while $\kappa(t)$ is a simple vibration of period $2\pi/\sqrt{\{-U_{qq}({}^0q)\}}$, if $U_{qq}({}^0q) < 0$.

This agrees with the last remark of §189 and explains why $F_{qq}({}^0q; {}^0h) \neq 0$ turned out to be negative in §189.

§191. The conditions of §187 characterize those solutions $q(t)$ of $[L]_q = 0$ which are periodic in the sense that, for some (but not for every) positive constant $\tau = \tau(h)$,

$$(12_1) \quad q(t + \tau) = q(t); \quad (12_2) \quad q'(t + \tau) = q'(t),$$

(12₂) being implied by (12₁). Since (12₂) does not imply (12₁), it remains to be seen whether or not one is justified in calling the solution $q(t)$ periodic in a case where only (12₂), i.e.,

$$(13) \quad q(t + \tau) = q(t) + \sigma,$$

is satisfied, where σ is independent of t and may or may not depend on the integration constants (or, what is the same thing, on the energy h of the solution). For instance, if q is an angular variable, to be reduced to a given modulus σ (e.g., $\sigma = 2\pi$ or $\sigma = 1$), it is unreasonable to define periodicity by (12₁) and not by (13) or, rather, by $q(t + \tau) \equiv q(t) \pmod{\sigma}$. And one can consider q as an angular variable which is to be reduced to modulus σ , if the coefficient functions $g(q)$, $U(q)$ of $L = \frac{1}{2}gq'^2 + U$ remain unchanged upon replacing q by $q + \sigma$, i.e., if the function $L(q', q)$ of q has for every fixed q' the period σ with respect to q . In fact, this condition characterizes those L for which $q = q(t) + \sigma$ is a solution of $[L]_q = 0$ for every solution $q = q(t)$.

§192. Now, if the functions $g(q)$, $U(q)$ of q have the period σ , the same holds for the function (1₂), where h is arbitrary. Suppose that the function (1₂) is, for a fixed h , positive for $-\infty < q < +\infty$. The periodicity condition of §187 is not satisfied and (3) is undefined, since α , β do not exist. However, if one defines τ by

$$(14) \quad \tau \equiv \tau(h) = \int_0^\sigma [F(q; h)]^{-\frac{1}{2}} dq,$$

it follows by an obvious modification of the uniqueness consideration of §187, that the solution of energy h is periodic in the sense (13) of angular periodicity.

§193. Suppose, for instance, that $L = \frac{1}{2}g(q)q'^2 + U(q)$ is given by $g(q) = 1$, $U(q) = \cos q$; so that $[L]_q \equiv q'' + \sin q = 0$ is the equation of motion of a pendulum in a Galilei field of gravitation, q being the angular distance from the vertical position. Then $\sigma = 2\pi$, while (1₂) reduces to* $F = 2(\cos q + h)$. If $h > 1$, the condition $F(q; h) > 0$, $-\infty < q < +\infty$, of §192 is satisfied; so that (13) is satisfied by (14) and $\sigma = 2\pi$, although $q = q(t)$ is, in view of (2), a steadily increasing or decreasing function which tends to $\pm\infty$ when either $t \rightarrow \pm\infty$ or $t \rightarrow \mp\infty$ (rotating pendulum). If $h = 1$, then $F = 4 \cos^2 \frac{1}{2}q$, so that the condition of §192 is not satisfied; while, $q_I = -\pi$ and $q_{II} = \pi$ being double roots of $F = 0$, the last remark of §186 shows that $q(t)$ tends to the pair of (mod 2π identical) equilibrium solutions $q = \pm\pi$, when $t \rightarrow \pm\infty$ (asymptotic movement towards the "unstable" vertical position of the pendulum). There is no solution of an energy $h < -1$, since $h < -1$ implies $F < 0$ for every q , which is impossible, by (1₁). If $h = -1$, then $F = -4 \sin^2 \frac{1}{2}q$; so that the solution is the equilibrium solution $q(t) \equiv 0$ (which represents the "stable" vertical position). If $-1 < h < 1$, then $F = 2(\cos q + h)$ does not satisfy the condition of §192 but it satisfies the conditions of §187 (oscillating pendulum with $\alpha \leq q \leq \beta$ ($= -\alpha < \pi$) as range of elongation).

If $h \rightarrow -1 + 0$, then $\beta = -\alpha \rightarrow +0$, and the period (3) tends, in view of (11), to 2π . This agrees with the last remark of §189, since the Jacobi equation belonging to the equilibrium solution $q(t) \equiv 0$ of $[\frac{1}{2}q'^2 + \cos q]_q \equiv q'' + \sin q = 0$ is $[\frac{1}{2}\kappa'^2 - \frac{1}{2}\kappa^2]_\kappa \equiv \kappa'' + \kappa = 0$, the equation of the pendulum with infinitesimal elongation. While in this case the characteristic exponents are $\pm i$ (hence, of the stable type; cf. §89), they become ± 1 (hence, of the unstable type) for the Jacobi equation $[\frac{1}{2}\kappa'^2 + \frac{1}{2}\kappa^2]_\kappa \equiv \kappa'' - \kappa = 0$ belonging to the equilibrium solution $q(t) \equiv \pm\pi$ which occurred in connection with the asymptotic case $h = 1$.

* Hence, the integral (2) is elliptic (of the first kind); and (3) is a complete elliptic integral.

Integrable Systems

§194. There will now be considered a class of problems with n degrees of freedom which are reducible to n problems with a single degree of freedom and are associated with the name of Liouville. These dynamical problems are characterized by the property that the $\frac{1}{2}n(n+1) + n + 1$ coefficient functions $g_{ik} = g_{ki}$, f_i , U of (1), §155 can be represented in terms of n sets of four functions $g_i(q_i)$, $f_i(q_i)$, $e_i(q_i)$, $d_i(q_i)$ of the single coordinate q_i , where $i = 1, \dots, n$, in the form $g_{ik} = \delta_{ik}g_i(q_i)G$, $f_i = f_i(q_i)$, $U = \sum e_i(q_i)/G$, where $G = \sum d_i(q_i)$; and $\delta_{ik} = 0$ for $i \neq k$, while $\delta_{ii} = 1$. Since $\sum f_i(q_i)dq_i$ is a complete differential, §156 shows that one may choose $f_i(q_i) \equiv 0$ without loss of generality; so that (1)-(2₁), §155 can be written as

$$(1_1) \quad L = \frac{1}{2}G \sum g_i q_i'^2 + G^{-1} \sum e_i; \quad (1_2) \quad G \equiv \sum d_i(q_i) > 0; \quad (1_3) \quad g_i(q_i) > 0.$$

Since the Hamiltonian function belonging to (1₁) is $H = \frac{1}{2}G^{-1} \sum g_i^{-1} p_i^2 - G^{-1} \sum e_i$ (§158), and since (1₂) satisfies the requirement of §180, the Hamiltonian function $\bar{H} = (-h + H)G$ of §180 becomes $\bar{H} = \sum H_i$, where $H_i = \frac{1}{2}g_i^{-1} p_i^2 - U_i$ and $U_i \equiv U_i(q_i; h) = e_i + h d_i$.

Since d_i , e_i , g_i depend only on q_i , the system $\dot{p}_i = -\bar{H}_{q_i}$, $\dot{q}_i = \bar{H}_{p_i}$ with n degrees of freedom can be replaced by the n systems with a single degree of freedom which one obtains by writing H_i for $\bar{H} = \sum H_i$. It is understood that the dots denote differentiations with respect to the time variable (14), §180, and that each of the n systems with a single degree of freedom has an energy integral $H_i = h_i$ in which one has to choose the n integration constants h_i so that $\sum h_i = 0$. In fact, the sum $\sum h_i$ of the partial energy constants is, in view of $\bar{H} = \sum H_i$, identical with the energy constant $\bar{H} = \bar{h}$ which, by the end of §180, must vanish.

Since the Lagrangian function belonging to $H_i \equiv H_i(p_i, q_i; h)$ is $L_i \equiv L_i(\dot{q}_i, q_i; h) = \frac{1}{2}g_i(q_i)\dot{q}_i^2 + U_i(q_i; h)$, one can apply §185-§186 to $[L_i]_{q_i} = 0$ for every i ; so that, in particular, $q_i = q_i(\bar{t})$ follows by the inversion of the quadrature which is assigned by the energy integral $\frac{1}{2}g_i\dot{q}_i^2 - U_i = h_i$. If there is known for every i a solution $q_i = q_i(\bar{t})$ of energy h_i , and if $\sum h_i = 0$, one sees from (16), §180 that the connection between \bar{t} and t is given by

$$(2_1) \quad t \equiv t(\bar{t}) = \sum s_i(\bar{t}); \quad (2_2) \quad s_i(\bar{t}) = \int d_i(q_i(\bar{t}))d\bar{t}; \quad \text{cf. } (1_2).$$

§195. Suppose, in particular, that, for certain fixed values of the

1 + n integration constants h, h_i which are subject to $\sum h_i = 0$, the conditions of §187 are satisfied for every i , it being understood that t is replaced by \bar{t} . Thus, the solution $q_i = q_i(\bar{t})$ of $[L_i]_{q_i} = 0$ has a period $\tau_i = \tau_i(h, h_i)$ with respect to \bar{t} and oscillates between an $\alpha_i = \alpha_i(h, h_i)$ and a $\beta_i = \beta_i(h, h_i)$. Consequently, (5)–(6), §188 hold if one replaces \bar{t} by a t_i , and t by \bar{t} , finally $G(q)$ by a corresponding $G_i(q_i)$, where $i = 1, \dots, n$. Accordingly, there are n time variables t_i , such that

$$(3) \quad 0 < \text{const.} < \dot{t}_i < \text{Const. for } -\infty < \bar{t} < +\infty \quad (\cdot = d/d\bar{t}),$$

where \bar{t} denotes the same time variable as in §194. Finally, from (7)–(9₂), §188,

$$(4_1) \quad q_i = \frac{1}{2}(\beta_i + \alpha_i) - \frac{1}{2}(\beta_i - \alpha_i) \cos t_i; \quad (4_2) \quad \bar{t} = t_i/\mu_i + r_i(t_i);$$

$$(4_3) \quad r_i(t_i + 2\pi) = r_i(t_i); \quad (4_4) \quad 0 < \tau_i = 2\pi/\mu_i.$$

According to (3), each of the n time variables $t_i = t_i(\bar{t})$ runs with \bar{t} from $-\infty$ to $+\infty$ in a monotone manner; while (4₁)–(4₄) imply that q_i , when considered as a function of \bar{t} , has the period τ_i .

Since $d_i = d_i(q_i)$, the function $d_i(q_i(\bar{t}))$ of \bar{t} also has the period τ_i . Let χ_i denote the constant term in the Fourier series of this periodic function; so that

$$(5_1) \quad d_i(q_i(\bar{t})) = \chi_i + c_i(\bar{t}); \quad (5_2) \quad c_i(\bar{t} + \tau_i) = c_i(\bar{t});$$

$$(5_3) \quad \chi_i = M\{d_i\},$$

where $M\{f\}$ denotes the limit of the mean value $\bar{t}^{-1} \int_0^{\bar{t}} f(t) dt$ of $f(\bar{t})$ when $\bar{t} \rightarrow \infty$ (so that $T^{-1} \int_0^T f(t) dt = M\{f\}$ in case $f(\bar{t})$ has the period T). It is clear from (5₁)–(5₃) and (2₂) that

$$(6_1) \quad s_i(\bar{t}) = \chi_i \bar{t} + v_i(\bar{t}); \quad (6_2) \quad v_i(\bar{t} + \tau_i) = v_i(\bar{t}).$$

Hence, (2₁) shows that $t = t(\bar{t})$ is the sum of the “secular” term $\chi \bar{t}$, where $\chi = \sum \chi_i = \text{const.}$, and of the “oscillating term” $\sum v_i(\bar{t})$, where $v_i(\bar{t})$ has the period τ_i .

§196. The τ_i are, in general, incommensurable, since every $\tau_i = \tau_i(h, h_i)$ is a continuous function of h, h_i . On the other hand, what one actually would like to have is the solution $q_i = q_i(t)$ of the n Lagrangian equations $[L]_{q_i} = 0$ belonging to the original Lagrangian function (1₁), where the independent variable is $t = \chi \bar{t} + \sum v_i(\bar{t})$. This requires an elimination of the $n + 1$ time variables t_i, \bar{t} between the $2n + 1$ parametrizations (4₁), (4₂), (2₁); an elimination which

clearly leads to n periodic functions $q_i = q_i(t)$ in the highly special case of n mutually commensurable \bar{t} -periods τ_i but involves, for unrestricted values of the time, a task of Diophantine intricacy in case at least two τ_i are incommensurable.

Nevertheless, one will expect that the $q_i(t)$ admit of an anharmonic Fourier analysis in case of arbitrary τ_i . In order to obtain this analysis, the non-local elimination of the $n + 1$ parameters t_i, \bar{t} will,* in §198, be carried out by using the theory of almost periodic functions ("almost periodicity" being meant in the sense of H. Bohr). The result will be that there exist n continuous functions $Q_i = Q_i(\vartheta_1, \dots, \vartheta_n)$ of n independent variables ϑ_i such that every Q_i has with respect to every ϑ_i the period 2π (i.e., every Q_i is a continuous function of the position on an n -dimensional $(\vartheta_1, \dots, \vartheta_n)$ -torus), and one has, for $-\infty < t < +\infty$ and $i = 1, \dots, n$,

$$(7) \quad q_i(t) = Q_i(\mu_1 t, \dots, \mu_n t), \quad \text{where} \quad \mu_i = 2\pi/\tau_i; \quad \text{cf. (4)}.$$

§197. If there exists between the n positive numbers μ_i a relation of the form $\sum N_i \mu_i = 0$, where the N_i are n integers such that $\sum N_i^2 \neq 0$, then one can (but need not) replace the dimension number, n , of the ϑ -torus by a smaller number. If n_0 denotes the least admissible value of the dimension number, then there exist exactly $n - n_0$ linearly independent relations $\sum N_i \mu_i = 0$, ($\sum N_i^2 \neq 0$), between the n positive numbers μ_i ; so that $n_0 = n$ in case the "frequencies" μ_i are "linearly independent," while $n_0 = 1$ in the trivial case where the partial periods $\tau_i = 2\pi/\mu_i$ are mutually commensurable.

Thus, if $n_0 = 1$, the solution path $q_i = q_i(t); i = 1, \dots, n$, is a closed curve in the n -dimensional configuration space; while if $n_0 = n$, one sees from (4₁)–(4₄) that, in virtue of what is called Kronecker's approximation theorem, the points (q_i) of the solution path $q_i = q_i(t); i = 1, \dots, n, -\infty < t < +\infty$, form a dense subset of the n -dimensional parallelepipedon $\alpha_i \leq q_i \leq \beta_i; i = 1, \dots, n$, in the configuration space. For the same reasons, the closure of the path is an n_0 -dimensional region† not only in the limiting cases

* The reader may omit the proof (i.e., §198), if he is not familiar with the theory of almost periodic functions.

† The situation is that while this fact depends only on Kronecker's approximation theorem, the Fourier analysis of the $q_i(t)$, i.e., the construction of the functions $Q_i(\vartheta_1, \dots, \vartheta_n)$ on a ϑ -torus, involves Weyl's refinement of Kronecker's theorem. Cf. the footnote to §127 bis.

$n_0 = 1$, $n_0 = n$ just mentioned but for any n_0 .

§198. The treatment of the problem of §196 will be based on two theorems concerning almost periodic functions. The first is the uniqueness theorem (on averages), while the second may be formulated as follows:

If $v = v(\bar{t})$, $-\infty < \bar{t} < \infty$, is a real-valued almost periodic function which has a derivative $\dot{v} \equiv dv/d\bar{t}$ satisfying an inequality of the form

$$(8) \quad -1 < -\theta < \dot{v}(\bar{t}), \text{ where } \theta = \text{const. and } -\infty < \bar{t} < +\infty,$$

and if the function $w = w(t)$, $-\infty < t < +\infty$, is defined by

$$(9) \quad \bar{t} \equiv t + w(t), \quad \text{where} \quad t \equiv \bar{t} + v(\bar{t}),$$

then the topological* mapping $t = \bar{t} + v(\bar{t})$ of the \bar{t} -axis upon the t -axis is such that the almost periodic function $t - \bar{t} \equiv v(\bar{t})$ of \bar{t} is an almost periodic function $t - \bar{t} \equiv -w(t)$ of t ; while the Fourier exponents of $v(\bar{t})$ and $w(t)$ determine the same modul.

In order to apply these facts, notice first that, by (1₂), the continuous function $G(q_1, \dots, q_n) = \sum d_i(q_i)$ has a positive minimum on the n -dimensional closed bounded region $\alpha_i \leq q_i \leq \beta_i$. Since $\alpha_i \leq q_i(\bar{t}) \leq \beta_i$, it follows that if $\text{fin inf } \sum d_i$ denotes the greatest lower bound of the function $\sum d_i(q_i(\bar{t}))$ for $-\infty < \bar{t} < +\infty$, then $\text{fin inf } \sum d_i$ is positive. Since the mean value, $M\{\sum d_i\}$, cannot be less than $\text{fin inf } \sum d_i$ and is, by (5₃), equal to $\sum \chi_i$, it follows that $\sum \chi_i > 0$; so that one can assume that $\sum \chi_i = 1$. In fact, this normalization involves only a change of the unit on the \bar{t} -axis, while introduction of a positive constant factor into the relation (2₁) which defines \bar{t} does not influence the preceding or following considerations. Thus,

$$(10) \quad 0 < \text{fin inf } \sum d_i(q_i(\bar{t})) \leq M\{\sum d_i\} = 1.$$

Put $v(\bar{t}) = \sum v_i(\bar{t})$; so that, from (2₁), (2₂) and (6₁),

$$(11_1) \quad t = \bar{t} + v(\bar{t}); \quad (11_2) \quad \dot{v}(\bar{t}) = -1 + \sum d_i(q_i(\bar{t})),$$

since $\sum \chi_i = 1$. Furthermore, (6₂) shows that $v(\bar{t}) = \sum v_i(\bar{t})$ is an almost periodic function, with frequencies which are contained in the

* Even if (8) is replaced by the weaker condition $-1 < \dot{v}(\bar{t})$, the function $t = \bar{t} + v(\bar{t})$ of \bar{t} is steadily increasing with \bar{t} from $-\infty$ to $+\infty$, since $v(\bar{t})$, being almost periodic, is bounded, while $\dot{t} = 1 + \dot{v} > 1 - 1 = 0$. However, $-1 < \dot{v}(\bar{t})$, $-\infty < \bar{t} < +\infty$, and the almost periodicity of $v(\bar{t})$ do not imply the almost periodicity of the function $w(t)$ which is uniquely defined by (9).

modul generated by the n (perhaps not linearly independent) numbers $\mu_i = 2\pi/\tau_i$. Now, two cases are possible, according as the alternative sign \leq in (10) is $<$ or $=$.

In the first case, (10) and (11₂) imply that (8) is satisfied by $\theta = 1 - \text{fin inf } \sum d_i$. Hence, the theorem mentioned in connection with (9) is applicable to (11₁), and so $\bar{t} = t + w(t)$, where $w(t)$ is an almost periodic function, with frequencies which are contained in the modul of the n numbers $\mu_i = 2\pi/\tau_i$. Consequently (7) follows from (4₁), if use is made of the representation (4₂)–(4₃) of $\bar{t} = t + w(t)$.

In the second case, (10) states that the greatest lower bound of $\sum d_i(q_i(\bar{t}))$ is identical with its mean value $M\{\sum d_i\} = 1$. Hence, the almost periodic function $\sum d_i(q_i(\bar{t}))$ is the constant 1. It follows, therefore, from (2₁)–(2₂) that $t = \bar{t}$ (up to an additive constant), and so (4₁)–(4₄) show that q_i , when considered as a function of t , is purely periodic for every i , with $\tau_i = 2\pi/\mu_i$ as period. Clearly, (7) holds in this degenerate case also.

§199. The result of §194 was that a system of the type (1₁)–(1₂) may be split into systems each of which has a single degree of freedom. The same situation occurs also when $n - 1$ of the n coordinates are ignorable (cf. §182–§184). Notice, however, that neither of these properties of a Lagrangian function is invariant under transformations of the configuration space or the phase space. For instance, if $n = 2$ and one replaces the Cartesian coordinates x, y by polar coordinates r, ϕ , it is quite possible that ϕ , but neither x nor y , is an ignorable coordinate (cf. §211). Correspondingly, while §117 implies that every dynamical system can be transformed, by means of a suitable canonical transformation, into a normal form (12)–(13), §113, in which all coordinates are ignorable, it is clear from the last remark of §113 that the main problem presents itself precisely in the construction of that suitable point transformation.

Actually, the situation is still less favorable. In fact, the proof of the existence of the suitable canonical transformations in question (or, what is the same thing, the existence proof for a complete solution W of (15), §114) can be based only on the general existence theorems of ordinary differential equations and implicit systems; theorems which are of a purely local nature by necessity. On the other hand, the actual mathematical questions of dynamics are not of this trivial local nature but present problems in the large which are controlled by the particular structure of the non-local topology

of the manifolds involved. This situation may be illustrated by a glance at the historical development of the idea of an “unsolved” dynamical problem.

§200. When John and James Bernoulli, Clairaut, D’Alembert, D. Bernoulli, Lambert, Euler and, finally, Lagrange applied the principles of Newton to the various problems of celestial and terrestrial mechanics, they had to face an awkward situation. For, on the one hand, it was almost axiomatic that a dynamical problem is “solved” only if it is reduced to quadratures (and successive differentiations and eliminations); while, on the other hand, the most urgent problems were almost never reducible to quadratures. The ingenious efforts of Clairaut ultimately led to a systematic theory of the lunar path and of the perturbations of the major planets, but not to the desired “solution by quadratures.”

Thus, it is understandable that Lambert became convinced that the problems of Celestial Mechanics may always be considered as “solved,” since, by means of numerical integrations of the equations of motion, these orbits can be calculated in advance with a high degree of numerical precision. From the beginning, the astronomers were compelled to develop, and be satisfied with, practicable procedures to this effect. During the following century, two of these numerical methods of astronomical origin, namely the “polygonal” method of finite differences and the method of successive approximations, became, in Cauchy’s hands, weapons of analytical existence or convergence proofs which, in turn, supplied a mathematical legalization of the numerical procedures of the astronomers. (The situation is similar in case of Newton’s method of undetermined coefficients, a method made legitimate by Cauchy’s principle of majorants.)

Since, on the one hand, these existence or convergence proofs have a general validity which has nothing to do with a dynamical problem, while, on the other hand, the simplest examples show that all these methods need be valid only on a restricted t -interval, everything that can be attained in this direction reduces to a manifestation of the local existence theorem of ordinary differential equations (cf. §79).

§201. From this point of view, a dynamical problem of which one knows its reducibility to quadratures but nothing more, can hardly be considered as being “solved” to a greater extent than a problem which is not reducible to quadratures. In fact, the quadratures in-

introduce functions which are not, in general, of an "elementary" type; so that, for actual computations or even only for qualitative information, recourse has to be made to mechanical quadratures (or, what is the same thing, to one of the approximating constructions, mentioned at the end of §200). Furthermore, what one usually wants are, not the functions represented by the quadratures, but rather the functions obtained by inversion of the system of quadratures (cf. §186); while the problem of inversion is, in general, a task which requires a machinery far more complicated than the existence theorem of ordinary differential equations (cf. §195–198).

These remarks imply that actually it is quite undefined what an "integrable" system is. It would be unnatural to make the notion of "integrability" of a dynamical system depend on the possibility of a reduction to quadratures. This is seen not only from §199 but also from examples which show that the possibility of a reduction to quadratures is neither sufficient nor necessary for a dynamical system which may be described by a sufficient degree of qualitative information (cf., on the one hand, §195–§198, and, on the other hand, the investigations concerning geodesics on two-dimensional manifolds with negative curvature, alluded to in §127). All of this lies along the line of Poincaré's dictum, according to which a system is neither integrable, nor non-integrable, but more or less integrable.

Concerning the present methodical situation, cf. §227 (and §440).

§202. In view of §185–§192, one will be inclined to consider a dynamical system as "integrable" if (but not only if) it can be split, by means of "explicit" transformations of the coordinates and the time variable, into a set of dynamical systems each of which has a single degree of freedom. In what follows, there will be considered a few classical cases of Lagrangian functions which satisfy these requirements.

§202 bis. As pointed out in §199, a Lagrangian function does not have the particular structure (1₁)–(1₂) in terms of arbitrary, but only in terms of suitably chosen, coordinates q_i .

For instance, Jacobi's result concerning the integrability of the problem of geodesics on a quadric $\alpha_1 x_1^2 + \alpha_2 x_2^2 + \alpha_3 x_3^2 = 1$ is to the effect that, if the three non-vanishing constants α_k are distinct* and if one applies elliptic coordinates as Gaussian parameters q_1, q_2 on

* Otherwise the surface is a surface of revolution, in which case the integration of the geodesic equations follows from §211.

the surface, then $ds^2 = dx_1^2 + dx_2^2 + dx_3^2$ appears in the form $ds^2 = G \cdot (g_1 dq_1^2 + g_2 dq_2^2)$, where g_i is a function of q_i alone and G is of the form (1₂); so that the Lagrangian function of the problem has, in terms of *these* coordinates q_1, q_2 , the structure (1₁)–(1₂), where $e_i(q_i) \equiv 0$ (incidentally, $G = q_2 - q_1$, while $g_i(q_i)$ is a quadratic rational expression in q_i). The same holds also when an $\alpha_k = 0$, in which case the elliptic coordinates degenerate into parabolic coordinates (cf. the end of §56).

Similarly, while the integration of two of the integrable cases of a top (namely, the case of a top without forces and the case of axial symmetry) can automatically be treated in terms of the underlying spherical coordinates (Euler and Lagrange), the integrability of the third integrable case is due to the fact that in this case (of Sonja Kowalewski) the Lagrangian function becomes of the type (1₁)–(1₂), if one introduces elliptic coordinates q_1, q_2 (Kolossoff).

§203. Consider the motion of a particle M under the Newtonian attraction of two bodies P_1, P_2 which are attracted neither by each other nor by M (Euler's problem of two fixed centra). Assume, for simplicity, that M moves in a plane; so that the problem has two, instead of three, degrees of freedom. Let x, y be the Cartesian coordinates of M . Choose the units of time, mass and distance so that the constant of gravitation, the sum of the masses of P_1 and P_2 , and the constant distance between P_1 and P_2 become unity. Furthermore, choose the origin and the orientation of the Cartesian coordinate system (x, y) so that $(0, 0)$ is the centre of mass of P_1 and P_2 , and that the fixed direction from P_1 towards P_2 is that of the positively oriented x -axis. Thus, if μ denotes the mass of P_2 , the mass of P_1 is $1 - \mu$; and P_1, P_2 rest at the points $(-\mu, 0), (1 - \mu, 0)$ of the (x, y) -plane. Consequently, if $x = x(t), y = y(t)$ denote the coordinates of M , and $r_1 = r_1(t), r_2 = r_2(t)$ the distances MP_1, MP_2 , then the Lagrangian function is

$$L = \frac{1}{2}(x'^2 + y'^2) + U, \text{ where } U = (1 - \mu)/r_1 + \mu/r_2.$$

Introduce instead of x, y the coordinates ξ, η of §56; so that

$$(12_1) \ 2r_1 = \cosh \eta + \cos \xi; \quad (12_2) \ 2r_2 = \cosh \eta - \cos \xi; \text{ cf. (34), §56.}$$

Then $\frac{1}{2}(x'^2 + y'^2) = \frac{1}{2}r_1r_2(\xi'^2 + \eta'^2)$, by what precedes (35), §56; so that

$$(13) \quad L = \frac{1}{2}r_1r_2(\xi'^2 + \eta'^2) + (r_1r_2)^{-1}\{(1 - \mu)r_2 + \mu r_1\}.$$

Substitution of (12₁)–(12₂) into (13) shows that L may be written in the form (1₁)–(1₂), §194, if one puts $n = 2$; $q_1 = \xi$, $q_2 = \eta$ and

$$g_1 = 1, g_2 = 1; \quad d_1 = -\frac{1}{4} \cos^2 \xi, d_2 = \frac{1}{4} \cosh^2 \eta; \\ e_1 = (\mu - \frac{1}{2}) \cos \xi, e_2 = \frac{1}{2} \cosh \eta.$$

Thus, the Lagrangian functions $L_i \equiv \frac{1}{2} g_i \dot{q}_i^2 + e_i + h d_i$ of §194 become $L_1 = \frac{1}{2} \dot{\xi}^2 + U_1$ and $L_2 = \frac{1}{2} \dot{\eta}^2 + U_2$, where

$$(14_1) \quad U_1 = \mu \cos \xi - \frac{1}{4} h \cos^2 \xi; \quad (14_2) \quad U_2 = \frac{1}{2} \cosh \eta + \frac{1}{4} h \cosh^2 \eta.$$

The energy integrals $\frac{1}{2} \dot{\xi}^2 - U_1 = h_1$, $\frac{1}{2} \dot{\eta}^2 - U_2 = h_2$ of the Lagrangian equations $[L_1]_{\xi} = 0$, $[L_2]_{\eta} = 0$ may be written as

$$(15_1) \quad \frac{1}{2} \dot{\xi}^2 - U_1 = h_0; \quad (15_2) \quad \frac{1}{2} \dot{\eta}^2 - U_2 = -h_0,$$

where h_0 is an arbitrary constant ($= h_1 = -h_2$, since $\sum h_i = 0$, by §194).

§204. Since (15₁), (15₂) are, by (14₁), (14₂), systems with one degree of freedom, §185–§188 are applicable† to ξ and η (§191–§192 only to ξ), it being understood that the dots denote differentiations with respect to the auxiliary time variable $\bar{t} = \bar{t}(t)$. Notice, however, that if the integration constants h , h_0 occurring in (14₁)–(14₂), (15₁)–(15₂) are chosen in a domain in which $\xi = \xi(\bar{t})$, $\eta = \eta(\bar{t})$ become periodic functions, and if $\tau_1 = \tau_1(h, h_0)$, $\tau_2 = \tau_2(h, h_0)$ denote the periods, then τ_1, τ_2 are continuous and non-constant functions of h, h_0 , and so not, in general, commensurable. Hence, unless $\tau_1 : \tau_2$ happens to be rational, the path $\xi = \xi(\bar{t})$, $\eta = \eta(\bar{t})$ of particle M under the attraction of P_1 and P_2 will not be periodic but such as to lie everywhere dense in a rectangle of the configuration plane (ξ, η) ; cf. §125.

§205. On proceeding in the same manner as in §193, one finds by a straightforward discussion of the (even) force functions (14₁)–(14₂), that the integration constants h, h_0 may be chosen in such a way that the periods τ_1, τ_2 are incommensurable, and that the closed (ξ, η) -rectangle on which the solution path is dense contains a point (ξ^*, η^*) at which $(\cos \xi^*, \cosh \eta^*) = (1, 1)$, but no point (ξ_*, η_*) at which $(\cos \xi_*, \cosh \eta_*) = (-1, 1)$. Since the rectangle is the closure of the set of those points to which the path $(\xi, \eta) = (\xi(\bar{t}), \eta(\bar{t}))$ of the particle M comes arbitrarily close as $\bar{t} \rightarrow \pm \infty$, it follows from (12₁)–(12₂) that $r_2 = r_2(\bar{t})$ does, and $r_1 = r_1(\bar{t})$ does not, come arbitrarily

† In view of (14_k), the quadrature assigned by (15_k) leads to an elliptic integral of the first kind; $k = 1, 2$.

close to zero for certain arbitrarily large values of \bar{t} . It is also seen from (12₁)–(12₂) that, at least for sufficiently large \bar{t} , one has $r_2(\bar{t}) > 0$ and not only $r_1(\bar{t}) > 0$. In fact, if r_2 vanished at certain values of \bar{t} which cluster at $\bar{t} = \infty$, the periods of the periodic functions $\cosh \eta(\bar{t})$, $\cos \xi(\bar{t})$ could not be incommensurable.

Now, $r_i = r_i(\bar{t})$ is the distance between the moving particle M and the fixed attracting centre P_i , where $i = 1, 2$. Hence, there is or is not a collision between M and P_i at a date \bar{t} according as $r_i(\bar{t}) = 0$ or $r_i(\bar{t}) > 0$. On the other hand, the choice of the integration constants just described leads to a motion of M such that both $r_i(\bar{t}) > 0$ for $\text{const.} < |t| < \infty$, although $\liminf r_2(\bar{t}) = 0$ as $\bar{t} \rightarrow \infty$. Consequently, the particle M can move under the attraction of the fixed centres P_1, P_2 in such a way that, although there is no actual collision between M and P_i , where $i = 1, 2$, the path of M penetrates an arbitrarily small circle about P_2 at certain arbitrarily distant dates \bar{t} .

Systems with Radial Symmetry

§206. If $n = 2$ and $L = \frac{1}{2}g(q_1)(q_1'^2 + q_2'^2) + U(q_1)$, one has to do with a particular case of (1₁)–(1₂), §194. This is seen by choosing $g_1 = g_2 = 1; d_2 = e_2 = 0, d_1 = g, e_1 = gU$.

As an example, consider the problem of geodesics on a surface S of revolution. Such a surface is characterized by the fact that, if one maps a domain on S upon a Euclidean (x, y) -plane in a suitable conformal way, and denotes by $g = g(x, y) > 0$ the factor of proportionality which, when multiplied by the Euclidean $dx^2 + dy^2$, gives the ds^2 on S , with x, y as Gaussian parameters on S , then $g(x, y)$ is a function of $(x^2 + y^2)^{\frac{1}{2}}$ alone. In other words, if r, ϕ , where

$$(1) \quad x = r \cos \phi, \quad y = r \sin \phi,$$

are chosen as Gaussian parameters on S , then the ds^2 on S becomes $ds^2 = g(r)(dr^2 + r^2d\phi^2)$, where $g(r) > 0$. Clearly, the equations of the meridians and of the parallel circles on S are $\phi = \text{const.}$ and $r = \text{Const.}$, respectively; while the geometrical meaning of $g(r)$ is that, if σ is the arc length on the meridian, then

$$(2) \quad d\sigma^2 = g(r)dr^2.$$

According to §178, the problem of geodesics on S is defined by the Lagrangian function

$$(3) \quad L = \frac{1}{2}s'^2, \quad \text{i.e.,} \quad L = \frac{1}{2}g(r)(r'^2 + r^2\phi'^2).$$

Hence, ϕ is an ignorable coordinate. Consequently, the Lagrangian equations admit, besides the energy integral, the integral $L_{\phi'} = \text{const.}$ According to (3), these integrals may be written as

$$(4_1) \quad \frac{1}{2}g(r)(r'^2 + r^2\phi'^2) = h; \quad (4_2) \quad g(r)r^2\phi' = c.$$

It is also seen from (3) that the problem reduces, for any fixed value of c , to that defined by the Lagrangian function $L^* = L^*(r', r; c)$ occurring in (22), §184, where $q_1 = r$, $g_{11} = g$, $g_{22} = r^2g$, $U = 0$; so that

$$(5) \quad L^* = \frac{1}{2}g^*(r)r'^2 + U^*(r; c), \text{ where } g^* = 1/g, \quad U^* = -\frac{1}{2}c^2g^*/r^2.$$

This is a problem $[L^*]_r = 0$, with a single degree of freedom, to which §185–§190 (and, if g^* , U^* are periodic functions of r , also §191–§192) are applicable. If a solution $r = r(t)$ of this reduced problem is known, $\phi = \phi(t)$ follows from (4₂) by a quadrature. In particular, $c = 0$ if and only if $\phi(t) = \text{const.}$, which means that the geodesic is a parallel circle, $r = r_0$.

§207. Consider the motion of a particle in an n -dimensional Euclidean space (x_i) under the action of a static central force, i.e., let

$$(6) \quad x_i'' = U_{x_i}(r); \quad i = 1, \dots, n, \text{ where } r = (x_1^2 + \dots + x_n^2)^{\frac{1}{2}}.$$

It will be shown that this problem is reducible to that of §206. Since this is obvious from §179 if $n = 2$ (cf. §212 below), it is sufficient to show that the case $n > 2$ is reducible to the case $n = 2$.

To this end, notice that if $j, k = 1, \dots, n$, then, from (6),

$$\begin{aligned} (x_j x_k' - x_j' x_k)' &\equiv x_j x_k'' - x_k x_j'' = x_j U_{x_k} - x_k U_{x_j} \\ &= (x_j x_k - x_k x_j) U_r / r \equiv 0. \end{aligned}$$

Hence, there exist integration constants c_{jk} for which

$$(7_1) \quad x_j x_k' - x_k x_j' = c_{jk}; \quad (7_2) \quad x_i c_{jk} + x_j c_{ki} + x_k c_{ij} = 0;$$

$$(7_3) \quad c_{jk} = -c_{kj} \quad (c_{ii} = 0),$$

(7₂), (7₃) being implied by (7₁) for arbitrary i, j, k ($= 1, \dots, n$).

It follows from (7₃) by a straightforward counting of the constants, that the set of all linear relations (7₂) determines a unique two-dimensional plane $\Pi = \Pi(c_{12}, \dots, c_{n-1, n})$ through the origin of the n -dimensional (x_i) -space, unless all $c_{jk} = 0$. Excluding, for a moment, the latter case, and noting that (7₁) represents integrals of (6),

and (7₂) dependencies between these integrals, one sees from the definition of an integral (§82), that if a solution path $x = x_i(t)$ of (6) belongs to the integration constants c_{jk} , then the path is contained in the plane Π , which is independent of t . Since (6) clearly is invariant under a (constant) rotation of the (x_i) -space about the origin $(x_i) = (0)$, one can choose the coordinate axes in such a way that Π becomes the (x_1, x_2) -plane. Then (6) holds for $n = 2$, while $x_i(t) \equiv 0$ for $i > 2$.

This completes the proof for the case in which not all $c_{jk} = 0$. In the remaining case, (7₁) shows that $x_j(t):x_k(t)$ is independent of t for all j, k , i.e., that the path $x_i = x_i(t)$ is contained in a line which is independent of t and goes through the origin of the (x_i) -space. Consequently, the plane Π exists also when all $c_{jk} = 0$, although Π is then not unique.

§208. As a consequence of §207, every conservative dynamical system which has radial symmetry and $n > 2$ degrees of freedom can be reduced, for every fixed value of the energy constant, to the problem treated in §206, where it is understood that the reductions involved require quadratures only.

First, the radial symmetry of a dynamical system defined by a Lagrangian function (1), §155 with n degrees of freedom is meant in the sense that (1), §155 remains invariant on arbitrary (constant) rotations of the n -dimensional Euclidean (q_i) -space about the origin $(q_i) = (0)$, if the coordinates q_i are chosen in a suitable manner. This clearly implies that in (1), §155 one has $(f_i) \equiv 0$ (up to a term of the type $(\sum q_i^2)'$, which may be omitted by §156), while U is a function of $(\sum q_i^2)^{\frac{1}{2}}$ alone; so that $\frac{1}{2}\sum \sum g_{ik}q_i'q_k'$ also is of radial symmetry. But it is known that a Riemannian space which carries a $ds^2 = \sum \sum g_{ik}dq_idq_k$ of radial symmetry can be mapped conformally on the Euclidean space, i.e. that, on replacing q_1, \dots, q_n by suitable new coordinates x_1, \dots, x_n , one has $ds^2 = g\sum dx_i^2$ for a suitable function g of proportionality; and that g and these coordinates x_i can be determined by mere quadratures and in such a way that g and $\sum q_i^2$ become functions of $\sum x_i^2$ alone.*

Consequently, $L = \frac{1}{2}g\sum x_i'^2 + U$, where g and U are functions of $r = (\sum x_i^2)^{\frac{1}{2}}$ alone. Finally, an application of the time transformation (14), §180 shows one can assume $g \equiv 1$ without loss of general-

* This fact is often used in the theory of relativity and can easily be proved by considering the geodesics which are transversal to a hypersurface $\sum q_i^2 = \text{const.}$

ity; so that $L = \frac{1}{2}\sum x_i'^2 + U(r)$. Since (6) belongs to this L , the proof is complete.

§209. The assumption $n > 2$ of §208 was necessary, since radial symmetry does not involve reversibility if $n = 2$. In fact, consider the system

$$x'' - 2\omega y' = U_x, \quad y'' + 2\omega x' = U_y,$$

where ω and U are given functions of (x, y) , and $\omega \equiv 0$ in the reversible case. It will be seen in §229 that this system has a Lagrangian function which, in virtue of (1), becomes

$$(8_1) \quad L = \frac{1}{2}(r'^2 + r^2\phi'^2) + r^2f(r)\phi' + U(r); \quad (8_2) \quad \omega(r) = \frac{1}{2}rf_r(r) + f(r),$$

if $\omega(x, y)$, $U(x, y)$ are functions of $r = (x^2 + y^2)^{\frac{1}{2}}$ alone. But (8₁) is of radial symmetry also in the irreversible case $\omega(r) \neq 0$, since the polar angle ϕ is an ignorable coordinate in (8₁).

For the latter reason, one has, besides the energy integral, the integral $L_{\phi'} = c$ ($= \text{const.}$), i.e., $r^2(f(r) + \phi') = c$.

§210. Let, in particular, $f(r) \equiv 1$ (so that (8₂) reduces to $\omega(r) \equiv 1$). Then $r^2(1 + \phi') = c$, i.e., $r^2\bar{\phi}' = c$, where $\bar{\phi} = t + \phi$. Furthermore, substitution of $f(r) \equiv 1$ and $\phi' = \bar{\phi}' - 1$ into (8₁) gives

$$(9_1) \quad L = \frac{1}{2}(r'^2 + r^2\phi'^2) + r^2\phi' + U(r);$$

$$(9_2) \quad L = \frac{1}{2}(r'^2 + r^2\bar{\phi}'^2) + \bar{U}(r), \quad \bar{U} = U - \frac{1}{2}r^2,$$

(9₁) and (9₂) being identical in virtue of $\phi = \bar{\phi} - t$ (cf. §95). Since (9₂) is and (9₁) is not of the reversible type, it follows that the notion of reversibility is not independent of the choice of the coordinate system.

This has, in the present case, a simple kinematical meaning. In fact, if $\bar{x} = r \cos \bar{\phi}$, $\bar{y} = r \sin \bar{\phi}$ and $x = r \cos \phi$, $y = r \sin \phi$, the identical Lagrangian functions (9₁), (9₂) become

$$(10_1) \quad L = \frac{1}{2}(x'^2 + y'^2) + (xy' - yx') + U;$$

$$(10_2) \quad L = \frac{1}{2}(\bar{x}'^2 + \bar{y}'^2) + \bar{U}; \quad (10_3) \quad U - \bar{U} = \frac{1}{2}r^2.$$

The transition from (\bar{x}, \bar{y}) to (x, y) represents the introduction of a Cartesian coordinate system (x, y) which rotates about the origin of (\bar{x}, \bar{y}) with constant angular velocity, since $\phi = \bar{\phi} - t$. That (10₂) is, and (10₁) is not, of the reversible type, is due to the Coriolis forces which are introduced by the rotation of the coordinate system (x, y) . Finally, the deviation, (10₃), of the force functions of (10₁)

and (10₂) is due to the centrifugal forces which are introduced by this rotation.

§211. Consider the motion of a particle in a Euclidean (x, y) -plane under the action of a force which is directed towards, or from, the origin $(x, y) = (0, 0)$, and has a magnitude $\pm F = |F|$ depending on the distance $r = (x^2 + y^2)^{1/2}$ only, where $F(r)$ is chosen as negative, positive or zero according as the force is, at the distance r , attractive, repulsive or neither. Then the equations of motion for the particle are given by $x'' = \pm F(r)x/r$, $y'' = \pm F(r)y/r$, or simply by (6), where $n = 2$, and $U(r)$ denotes the undetermined integral of $\pm F(r)$. Thus, $U = U(\sqrt{x^2 + y^2})$ and

$$(11_1) \quad x'' = U_x, \quad y'' = U_y;$$

$$(11_2) \quad \frac{1}{2}(x'^2 + y'^2) - U = h; \quad (11_3) \quad xy' - yx' = c,$$

(11₂), (11₃) being integrals of (11₁). Introducing polar coordinates, one has

$$(12_1) \quad L = \frac{1}{2}(r'^2 + r^2\phi'^2) + U(r);$$

$$(12_2) \quad \frac{1}{2}(r'^2 + r^2\phi'^2) - U(r) = h; \quad (12_3) \quad r^2\phi' = c,$$

since (12₂), (12₃) and (12₁) are, in virtue of (1), identical with (11₂), (11₃) and the Lagrangian function $L = \frac{1}{2}(x'^2 + y'^2) + U$ of (11₁). It is seen from (12₁) that the angle ϕ is an ignorable coordinate, and that the momentum $L_{\phi'}$ canonically conjugate to this angular coordinate is $r^2\phi'$. For this reason, the integral (12₃), i.e. (11₃), is usually referred to as expressing the conservation of angular momentum; while (12₂), i.e. (11₂), represents the conservation of energy.

§212. If one excludes the trivial case of an equilibrium solution, as well as the isolated t which belong to cusps, §179 shows that those solutions of (11₁) which have the energy h can be interpreted as the geodesics on the surface S_h on which the square of the line element is the product of g and $dx^2 + dy^2$, where g is the function $2(U + h)$ of the Gaussian parameters x, y . Hence, $g = g(r)$, and so S_h is, for every fixed h , a surface of revolution, considered in §206. Since $U = U(r)$, where $r^2 = x^2 + y^2$, the Gaussian curvature $K_h = K_h(x, y)$ on S_h is readily found to be

$$(13) \quad K_h \equiv K_h(r) = \frac{1}{4} \{ U_r^2 - (U + h)(U_{rr} + U_r/r) \} / (U + h)^3;$$

(cf. (19), §231). For instance, the metric on S_h becomes non-Euclidean if

(14) $U = 2(1 - r^2)^{-2}$ and $h = 0$, since then $K_h(r) \equiv -1$, by (13).

§213. Since there exist for every solution of (11₁) constants h, c which satisfy (11₂), (11₃), one might expect that the condition which is imposed by the pair of conditions (11₁) on a pair of functions $x = x(t), y = y(t)$ of class $C^{(2)}$ is equivalent to the pair of conditions (11₂), (11₃), the values of the constants h, c being unspecified. Actually, the necessary conditions (11₂)–(11₃) for (11₁) are sufficient as well if one excludes the case of a circular solution. But if $x(t)^2 + y(t)^2 = \text{const.}$, then (11₂)–(11₃) do not imply (11₁).

In fact, differentiation of (11₂)–(11₃) gives

$$x'x'' + y'y'' - U_x x' - U_y y' = 0, \quad xy'' - yx'' = 0;$$

so that, since $yU_x - xU_y \equiv 0$ in view of $U = U(\sqrt{x^2 + y^2})$, the pair (11₂)–(11₃) is equivalent to the pair

$$(15) \quad x'(x'' - U_x) + y'(y'' - U_y) = 0, \quad y(x'' - U_x) - x(y'' - U_y) = 0.$$

And the equations (15) are linear combinations of the equations (11₁), with $-xx' - yy' \equiv -\frac{1}{2}(x^2 + y^2)'$ as determinant; so that (15) and (11₁), i.e. (11₂)–(11₃) and (11₁), are equivalent, unless $x(t)^2 + y(t)^2 = \text{const.}$

§214. According to (12₁) and §184, one can replace (11₁), for every fixed value of the constant (11₃), by $[L^*]_r = 0$, where

$$(16_1) \quad L^* = \frac{1}{2}r'^2 + U^*; \quad (16_2) \quad U^*(r; c) = U(r) - \frac{1}{2}c^2/r^2;$$

$$(16_3) \quad \frac{1}{2}r'^2 - U^* = h.$$

It is clear from (12₂), (12₃), (16₂) that (16₃) represents not only the energy integral of the system $[L^*]_r \equiv r'' - U_r^* = 0$ with a single degree of freedom but also the energy integral, (11₂), of the system (11₁) with two degrees of freedom. If a solution $r = r(t)$ of $[L^*]_r = 0$ is known, then $\phi = \phi(t)$ in (1) follows from (12₃) by a quadrature.

It is also seen from (12₃) that the path in the (x, y) -plane is direct or retrograde for every t according as $c > 0$ or $c < 0$, and that, if one changes t to $-t$, a retrograde path becomes direct. Since (11₁) is of the reversible type (§156), it follows that the path can be assumed to be either direct or such that $c = 0$. Finally, (11₃) shows that $c = 0$ if and only if the path in the (x, y) -plane is contained in a fixed line through the origin.

§215. If a solution $r = r(t)$ of $[L^*]_r \equiv r'' - U_r^* = 0$ is such that neither $c = 0$ nor $r(t) \equiv r_0$, where $r_0 = \text{const.}$, then $r(t)$ is, by §185–§187, either of the asymptotic type or such as to have the period τ , where

$$(17_1) \quad \tau = 2 \int_{\alpha}^{\beta} [2(U^*(r; c) + h)]^{-\frac{1}{2}} dr;$$

$$(17_2) \quad \alpha = \min r(t) < \max r(t) = \beta.$$

Consider the latter case and assume that $r(t) \neq 0$ for every t , i.e., that $\alpha > 0$. Denoting by $\gamma > 0$ the constant term in the Fourier series of the continuous periodic function $1/r(t)^2$, and placing $\nu = c\gamma$, one sees from (12₃) that $\phi(t) = \nu t + \psi(t)$, where $\psi(t)$ has the same period, τ , as $r(t)$. Hence, it is clear from (1) that the qualitative behavior of the path in the (x, y) -plane as $t \rightarrow \infty$ depends on whether the value of the integration constant $\nu\tau:\pi$ is rational or irrational. In fact, in the first case both functions $x(t)$, $y(t)$ have a multiple of τ as period; so that the path in the (x, y) -plane closes into itself after a sufficient number of circuits. If, on the other hand, $\nu\tau$ and π are incommensurable, it is clear from the corresponding remarks† of §125 or §197, that the path in the (x, y) -plane comes, as $t \rightarrow \infty$, arbitrarily close to every point of the circular ring $\alpha^2 \leq x^2 + y^2 \leq \beta^2$, where α, β are defined by (17₂).

§216. Suppose that $r(t) \equiv r_0$, where $r_0 = \text{const.} > 0$, is an equilibrium solution of $[L^*]_r \equiv r'' - U_r^* = 0$, and denote by c_0, h_0 the constants c, h which belong to this solution; so that, from (16₂)–(16₃),

$$(18_1) \quad -c_0^2 = r_0^3 U_r(r_0), \quad -h_0 = U(r_0) + \frac{1}{2} r_0 U_r(r_0);$$

$$(18_2) \quad 2r_0^2(U(r_0) + h_0) = c_0^2,$$

(18₂) being implied by the pair of conditions (18₁) which are necessary and sufficient for the existence of a number $r_0 > 0$ such that $r(t) \equiv r_0$ is an equilibrium solution belonging to the values of $c_0 \geq 0$ and $h_0 \leq 0$ defined by (18₁). In view of (18₁), one can start with an arbitrary $r_0 > 0$ if and only if $U_r(r) < 0$ for every r , which means, by §211, that the force is attractive at every r .

It is clear from (1) and (12₃) that an equilibrium solution $r(t)$

† Actually, the configuration region on which the path is dense must again be thought of as a torus, since $r = r(t)$ is periodic and $\phi = \phi(t)$ has to be reduced mod 2π .

$\equiv r_0 (> 0)$ of $[L^*]_r = 0$ represents an equilibrium solution or a circular solution of (11₁) according as $c_0 = 0$ or $c_0 > 0$, and that in the latter case the angular velocity of the motion along the circle $x^2 + y^2 = r_0^2$ has the constant value c_0/r_0^2 ; so that, on denoting the period $2\pi r_0^2/c_0$ by τ_0 and using (18₁), one has

$$(19_1) \quad x = r_0 \cos 2\pi t/\tau_0, \quad y = r_0 \sin 2\pi t/\tau_0;$$

$$(19_2) \quad \tau_0 = 2\pi : \sqrt{\{ -U_r(r_0)/r_0 \}}.$$

Thus, the circular motions are periodic, whether a commensurability condition (cf. §215) is satisfied or not.†

According to §190, the characteristic exponents of the equilibrium solution $r(t) \equiv r_0$ of $[L^*]_r \equiv r'' - U_r^*(r; c_0) = 0$ are the square roots of $U_{rr}^*(r_0; c_0)$ and can, therefore, be written as

$$(20) \quad \pm \{ U_{rr}(r_0) + 3U_r(r_0)/r_0 \}^{\frac{1}{2}},$$

since $U_{rr}^*(r_0; c_0) = U_{rr}(r_0) - 3c_0^2/r_0^4$, $c_0^2 = -r_0^3 U_r(r_0)$, by (16₂), (18₁).

§217. In what follows, it will be assumed that the circular solution (19₁) exists for every given $r_0 > 0$. According to §216, this will be the case if and only if $U_r(r)$ is negative for every r .

It is natural to ask when has the law $U_r(r)$ of attraction the property that every solution $x = x(t)$, $y = y(t)$ whose integration constants h, c are sufficiently close to the integration constants h_0, c_0 of a circular solution (19₁) is a periodic solution of (11₁) and has a period $\tau = \tau(c, h)$ which tends to the corresponding circular period (19₂), as $c \rightarrow c_0$, $h \rightarrow h_0$. In view of the Diophantine situation described at the end of §215, it is not surprising that the restriction imposed by this assumption is so heavy as to make possible an explicit determination of the function $U_r(r)$. In fact, it will turn out that, with the exception of a trivial case, the force $U_r(r)$ must be proportional to r^{-2} ; so that the law of attraction is Newtonian (as to the trivial exception, cf. §219 bis).

§218. In order to prove this, choose an unspecified pair of constants $c_0 > 0, h_0$ which satisfy the conditions (18₁) for a circular solution of suitable radius r_0 , and, keeping c_0 fixed, let the integration constant (11₂) of (11₁) vary close to the fixed h_0 in an arbitrary way. Then, by the requirement of §217, the radius vector $r(t)$ of the solu-

† It was tacitly assumed in §215 that the periodic remainder term $\psi(t)$ of $\phi(t) = \nu t + \psi(t)$ is not independent of t . But if it is, then $\phi'(t) = \nu = \text{const.}$ And this leads, by (12₃), to the circular case $r(t) = \text{Const.}$, excluded in §215.

tion of (11₁) which belongs to the integration constants c_0 , h must have a period $\tau = \tau(c_0; h)$. Also

$$(21) \quad \lim_{h \rightarrow h_0} \tau(c_0; h) = 2\pi: \{ -U_{rr}(r_0) - 3U_r(r_0)/r_0 \}^{\frac{1}{2}} > 0.$$

In fact, the characteristic exponents of $r(t) \equiv r_0$ are given by (20); so that (21) is implied by §189–§190.

Since $r_0 > 0$ is arbitrary by the assumption of §217, it follows from (21) that, for every $r_0 > 0$,

$$(22) \quad U_{rr}(r_0) + 3U_r(r_0)/r_0 < 0; \quad \text{and} \quad U_r(r_0) < 0, \quad \text{by §217.}$$

But §217 requires also that the limit (21) be equal to the circular period (19₂) for every $r_0 > 0$. Hence, on writing r instead of r_0 , one must have

$$(23) \quad U_{rr}(r) + 2U_r(r)/r = 0.$$

Now, the general solution of the linear differential equation (23) for $U = U(r)$ is Const./ r plus a constant. Since only the force $U_r(r)$ is of interest, one can choose this additive constant to be 0. And (22) requires that Const. > 0 ; so that Const. = 1 upon a suitable choice of the unit of length. Thus, $U(r) = 1/r$, as stated in §217.

It will turn out in §241 (and §267) that this particular U actually satisfies the requirements of §217.

§218 bis. While the law $U(r) = r^{-1}$ has thus been obtained by a consideration of nearly circular orbits, it is important to realize the exceptional rôle of this law from the general point of view of §126–§130. In this regard, it will turn out in §241 that, in the case $U = r^{-1}$, there exists for every solution $x = x(t)$, $y = y(t)$, nearly circular or not, two integration constants a, b such that $cx' = -yU + a$, $cy' = xU + b$ for every t . Hence, if $U = r^{-1}$, the single integral (11₂) may be replaced by the two integrals

$$(11 \text{ bis}) \quad \begin{aligned} (xy' - yx')x' + yU &= a, \\ (xy' - yx')y' - xU &= b, \quad \text{where} \quad U = (x^2 + y^2)^{-\frac{1}{2}}; \end{aligned}$$

so that, instead of the two integrals (11₂)–(11₃) of (11₁), one has three conservative integrals which are independent in the sense of §82 and, being algebraic, represent isolating integrals in the sense of §128. If, on the other hand, $U = U(r)$ is arbitrary, one has for (11₁) only

the two isolating integrals (11₂)–(11₃); while the third conservative integral (which exists by §82, and depends, by §214, on the inversion of a quadrature) is not of the isolating type, the reason being sufficiently clear from §215. Thus, for a general $U(r)$, the degree $m - 1 - l$ of primitivity (§130) is $4 - 1 - 2 = 1 \neq 0$, but it reduces to $4 - 1 - 3 = 0$ in Newton's case. It will be seen from §219–§219 bis that this reduction takes place in Hooke's case also, but in no case distinct from these two.

§219. The condition imposed in §217 on U requires not only that every solution path which is close to a circular solution be periodic but also that the period of such a solution be close to the corresponding circular period, (19₂). It is natural to ask, which $U(r)$ are obtained if this additional restriction is omitted. Then one has to allow that the limit (21), instead of being equal to the circular period (19₂), is only commensurable with it; so that the circular limit of a nearly circular orbit is thought of as returning into itself after an unspecified number of circuits. Thus, comparison of (19₂) with (21) shows that (23) must be replaced by its generalization

$$(23 \text{ bis}) \quad U_{rr}(r) + (3 - \lambda^2)U_r(r)/r = 0,$$

where λ is some fixed rational number. The general solution of (23 bis) for any fixed λ is seen to be $U(r) = \text{const.} \cdot r^{\lambda^2-2}$ (plus a constant which may, as in §218, be chosen to be 0).

However, it turns out that the value $\lambda = 1$, found in §218, is the only admissible value of λ . In fact, if λ is chosen to be distinct from 1, then a detailed discussion of (12₂)–(12₃) for $U = \text{const.} \cdot r^{\lambda^2-2}$ shows that the solutions close to a circular solution cannot be all periodic (except in the case $\lambda = 2$, which can be ruled out for another reason*). This is not surprising, since (23 bis) was derived only from the necessary condition (21), which does not happen to be sufficient as well. In fact, in order to prove that every $\lambda \neq 1$ must be excluded, one has to calculate for the hypothetical period $\tau = \tau(c_0, h)$, where $|h_0 - h|$ is small, an approximation which is by one degree higher than the 0-th approximation supplied by (21). This elementary, though somewhat lengthy, calculation will not be carried out here.

* It will be seen in §219 bis that if $\lambda = 2$, the solution paths in the (x, y) -plane are ellipses, hence *simple* closed curves, and belong therefore to $\lambda = 1$, by the definition of λ in (23 bis); so that the assumption $\lambda = 2$ leads to the contradiction $2 = 1$.

§219 bis. There exists a singular case, which was neglected above. In fact, it is clear from §189–§190 that the considerations of §218 break down in case U is such as to make the period independent of the integration constants. In view of (19₂), §216, this assumption of tautochronism supplies for $U = U(r)$ the condition $U_r(r)/r = \text{const.}$; while the requirement $U_r < 0$ of §217 shows that $\text{const.} < 0$, and so $\text{const.} = -1$ without loss of generality. Thus, $U_r = -r$, i.e., $U = -\frac{1}{2}r^2$ (plus a superfluous constant). Hence, (11₁) reduces to $x'' + x = 0$, $y'' + y = 0$, and has, therefore, the general solution $x = a \cos(t - t^0)$, $y = b \cos(t - t_0)$, which is always periodic (the equilibrium solution ($a = 0 = b$) is excluded by the assumption $r > 0$). The fact that the period ($= 2\pi$) is independent of the integration constants agrees with the end of §160 bis, since $\beta^{-1} - \frac{1}{2} = 0$ in the present case. Clearly, one has to do with the subcase $\omega_1 = \omega_2$ of the case (i) of §125 (cf. §130).

It will turn out in §259 that the case $U = r^{-1}$ of §217–§218 is, for fixed $h (< 0)$, reducible to the present trivial case $U = -\frac{1}{2}r^2$.

§220. Consider again the general case of §211. Exclude, for simplicity, the exceptional cases mentioned at the beginning of §212. Let Φ, R denote the momenta $L_{\phi'}$, $L_{r'}$ canonically conjugate to ϕ, r . Thus, $\Phi = r^2\phi'$, $R = r'$, by (12₁). Hence, the Hamiltonian function $H(\Phi, R; \phi, r)$ belonging to (12₁) is

$$(24) \quad \begin{aligned} H &\equiv H(\Phi, R; r) \\ &= \frac{1}{2}(R^2 + \Phi^2/r^2) - U(r), \text{ by (2}_1\text{), §15; so that } H = h, \Phi = c \end{aligned}$$

are the integrals (12₂), (12₃) of energy and of angular momentum. The partial differential equation (15), §114 becomes

$$(25) \quad \frac{1}{2}(W_r^2 + W_\phi^2/r^2) - U(r) = h, \text{ where } (\phi, r) \equiv (q_1, q_2) \equiv q.$$

§221. Notice that $q_1 = \phi$ does not occur explicitly in (25); while $W_\phi \equiv \Phi = c$, by §220. This, when compared with the passage (16), §114 from (14₂), §114 to (15), §114, suggests for (25) the existence of a solution $W = W(\phi, r)$ of the type $W = c\phi + V$, where $V = V(r)$ is independent of ϕ . Actually, the partial differential equation (25) then reduces to $\frac{1}{2}(V_r^2 + c^2/r^2) - U(r) = h$. And this is an ordinary differential equation whose general solution $V = V(r)$ follows by a quadrature as the undetermined integral of $\{2(U(r) + h) - c^2/r^2\}^{\frac{1}{2}}$. Thus, if $r^0 = r^0(c, h)$ denotes an unspecified function of the integration constants c, h , then

$$(26) \quad W = c\phi + V \equiv c\phi + \int_{r^0(c,h)}^r \{2(U(\bar{r}) + h) - c^2/\bar{r}^2\}^{\frac{1}{2}} d\bar{r}$$

is a solution $W = W(\phi, r)$ of (25). However, caution is necessary, since $W(\phi, r)$ must possess continuous derivatives W_r, \dots . And this condition is violated if the integrand $\{ \ }^{\frac{1}{2}}$ of (26) vanishes. But comparison of (26) with (18₂) shows that $\{ \ }^{\frac{1}{2}}$ vanishes identically precisely in case of a circular solution. On the other hand, the vanishing of the integrand $\{ \ }^{\frac{1}{2}}$ at an isolated r of the path (say, at $r = r^0$) does not matter, of course.

§222. Barring the case of circular solutions, (26) represents a complete solution of (25), the integration constants v_1, v_2 of §116 being represented by c, h . In fact, the completeness condition (18), §116 is then satisfied, since

$$\det (W_{q_i v_k}) \equiv \begin{vmatrix} W_{\phi c} & W_{\phi h} \\ W_{rc} & W_{rh} \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ W_{rc} & V_{rh} \end{vmatrix} = V_{rh} = \frac{1}{\{ \ }^{\frac{1}{2}}}, \text{ by (26).}$$

Hence, the rule of §116 bis is applicable to $q_1 = \phi, q_2 = r; Q_1 = c, Q_2 = h$. But $-W_{Q_1} \equiv -W_c = -\phi - V_c$, by (26). Hence, if t^0 is a fixed date, and f^0 denotes $(f)_{t=t^0}$ for any f , the rule of §116 bis shows that

$$(27) \quad P_1 = -\phi^0 - V_c^0, \quad P_2 = t^0; \quad Q_1 = c, \quad Q_2 = h$$

is a canonical set of integration constants.

§223. Let the derivative $r' = r'(t)$ of the solution $r = r(t) \equiv r(t; c, h) > 0$ of (16₃) vanish at some isolated $t = t_0 = t_0(c, h)$; for instance, let $r(t_0)$ be a local minimum of $r(t)$. Suppose that the t^0 of §222 is chosen to be this t_0 , and let the unspecified lower limit $r^0 = r^0(c, h)$ of the quadrature (26) be identified with $r(t_0)$. Then, under obvious assumptions of differentiability,

$$(28) \quad P_1 = h, \quad P_2 = c; \quad Q_1 = -t_0, \quad Q_2 = \omega, \text{ where } \omega = \phi(t_0), (r'(t_0) = 0),$$

is a canonical set of integration constants.

This may be proved with the use of the concluding remark of §221. First, from (26),

$$V_c \equiv \left(\int_{r^0(c,h)}^r \{ \ }^{\frac{1}{2}} d\bar{r} \right)_c \equiv -r_c^0(c, h) (\{ \ }^{\frac{1}{2}})^0 + \int_{r^0(c,h)}^r (\{ \ }^{\frac{1}{2}})_c d\bar{r}.$$

Since the last integral vanishes at $r = r^0(c, h) \equiv r(t_0)$, it follows that

$V_c^0 = -r_c^0(c, h)(\{ \}^{\frac{1}{2}})^0$. But $(\{ \}^{\frac{1}{2}})^0 = 0$, as seen by placing $t = t^0 \equiv t_0$ in (16₂)–(16₃), and then using the assumption $r'(t_0) = 0$. Consequently, $V_c^0 = 0$. Hence, (27) reduces to a set of canonical integration constants which is, in view of §42, equivalent to (28).

§224. These results will now be transferred to the case of three Cartesian coordinates. To this end, consider, for a fixed value of a parameter ι , that conservative transformation $v = v(q)$ of $n = 3$ coordinate variables $q_1 = x$, $q_2 = y$, $q_3 = \nu$ into coordinate variables $v_1 = \xi$, $v_2 = \eta$, $v_3 = \zeta$ which is given by

$$(29) \quad \begin{aligned} \xi &= x \cos \nu - y \sin \nu \cos \iota, & \eta &= x \sin \nu + y \cos \nu \cos \iota, \\ \zeta &= y \sin \iota. \end{aligned}$$

The Jacobian matrix $J \equiv v_q$ of (29) is seen to be

$$(30) \quad J = \begin{pmatrix} \cos \nu & -\sin \nu \cos \iota & -\eta \\ \sin \nu & \cos \nu \cos \iota & \xi \\ 0 & \sin \iota & 0 \end{pmatrix},$$

the i -th row of (30) representing the partial derivatives of v_i with respect to x, y, ν . It will be assumed that

$$(31_1) \quad -\sin \iota \neq 0 \quad (\text{i.e., } \iota \neq 0, \pm \pi, \dots);$$

$$(31_2) \quad \xi \cos \nu + \eta \sin \nu \neq 0,$$

since the determinant of (30) is the product of (31₁) and (31₂). Placing

$$(32) \quad \begin{aligned} X &= \Xi \cos \nu + H \sin \nu, \\ Y &= (-\Xi \sin \nu + H \cos \nu) \cos \iota + Z \sin \iota, \\ N &= -\Xi \eta + H \xi, \end{aligned}$$

one sees that Ξ, H, Z are momenta canonically conjugate to the coordinates ξ, η, ζ , if X, Y, N are momenta canonically conjugate to x, y, ν . In fact, (6), §49 states that the canonical extension of the present coordinate transformation (29) is obtained by transforming the 3-vector (X, Y, N) into (Ξ, H, Z) by the matrix J^{-1} ; so that J transforms (Ξ, H, Z) into (X, Y, N) . Now, (30) shows that the matrix of the substitution (32) is actually J .

Accordingly, (29) and (32) together represent, for every fixed $\iota = \text{const.}$ satisfying (31₁), a completely canonical transformation of

$(X, Y, N; x, y, \nu)$ into $(\Xi, H, Z; \xi, \eta, \zeta)$, provided that (31₂) is satisfied.

In order to interpret (29), notice first that if the x -axis lies in the (ξ, η) -plane of Fig. 1 (§78), i.e., if $\omega = 0$, then (24), §78 reduces to (23), §78. But, on transforming a 3-vector (x, y, z) into a 3-vector by the rotation (23), §78, and then placing $z \equiv 0$, one obtains precisely the above transformation (29).

Accordingly, an interpretation of (29) is that a particle moving in an (x, y) -plane is referred to a coordinate system (ξ, η, ζ) whose (ξ, η) -plane contains the x -axis. And ι denotes the "inclination" of the (x, y) -plane towards the (ξ, η) -plane, while ν , the "node," is the angular distance of the x -axis from the ξ -axis; cf. Fig. 1 (§78).

§225. Let (11₁) be replaced by

$$(33) \quad \begin{aligned} \xi'' &= U_\xi(r), & \eta'' &= U_\eta(r), & \zeta'' &= U_\zeta(r), & \text{where} \\ r &= (\xi^2 + \eta^2 + \zeta^2)^{\frac{1}{2}}. \end{aligned}$$

Accordingly to §207, any solution path of (33) lies in a plane through the origin of the (ξ, η, ζ) -space, and is a rectilinear path only when all the integration constants (7₁) vanish. Consider a fixed solution path which is not rectilinear, and choose its plane as an (x, y) -plane whose x -axis lies within the (ξ, η) -plane. Then the situation is that described at the end of §224; so that (29) is valid. Since the coordinate ν and the parameter ι are independent of t , differentiation of (29) gives

$$(34) \quad \begin{aligned} \xi' &= x' \cos \nu - y' \sin \nu \cos \iota, & \eta' &= x' \sin \nu + y' \cos \nu \cos \iota, \\ \zeta' &= y' \sin \iota. \end{aligned}$$

It will be assumed that (31₁), (31₂) are satisfied.

The Hamiltonian function belonging to the Lagrangian equations (33) clearly is $H = \frac{1}{2}(\Xi^2 + H^2 + Z^2) - U(r)$, with $\Xi = \xi'$, $H = \eta'$, $Z = \zeta'$ as momenta canonically conjugate to the coordinates ξ, η, ζ . Comparing this with (32), (34) and (29), one readily verifies that $X = x'$, $Y = y'$, $N = (-x'y + y'x) \cos \iota$. But $(-x'y + y'x) = c$, by (11₃). Hence, the momenta X, Y, N canonically conjugate to the coordinates x, y, ν are $x', y', c \cos \iota$, where the inclination ι of the path is considered as fixed (cf. the beginning of §224). Furthermore, the canonical transformation involved is completely canonical, since so is the transformation considered in §224.

§226. To the Lagrangian passage from (11₁) to (12₁) there corresponds the Hamiltonian passage from the coordinates x, y and momenta $X = x', Y = y'$ to the coordinates ϕ, r and momenta $\Phi = r^2\phi', R = r'$, considered in §220. According to §49, this Hamiltonian passage represents a completely canonical transformation, since it is seen to be the canonical extension of (1). On the other hand, application of §221–§222 to the case of §223 has shown that the passage from $\Phi, R; \phi, r$ to (28) is canonical and of multiplier $\mu = 1$, since this property belongs to the definition (§104) of a canonical set of integration constants. On combining these facts with the result of §225, one sees that the passage from the momenta ξ', η', ζ' and coordinates ξ, η, ζ of (33) to

$$(35) \quad p_1 = h, \quad p_2 = c, \quad p_3 = c \cos \iota; \quad q_1 = -t_0, \quad q_2 = \omega, \quad q_3 = \nu$$

$$(\omega = \phi(t_0), \quad r'(t_0) = 0)$$

is a canonical transformation of multiplier $\mu = 1$. And these p_i, q_i are, by §225, independent of t .

Consequently, (35)' represents a canonical set of integration constants of (33).

This is the result indicated at the beginning of §224. It should be mentioned that the canonical integration constants (35) could have been obtained, without the use of (34), by using §224 only. This direct procedure would have been more explicit, but also more lengthy, than the way followed above.

Two Degrees of Freedom

§227. A conservative dynamical system with $n = 2$ degrees of freedom is not, in general, "integrable." Furthermore, only a few of these non-integrable systems have thus far been studied in any detail. Finally, it is quite possible that these particular non-integrable systems belonging to $n = 2$ are not involved enough to present the characteristic difficulties which might arise in the "generic" case $n = 2$.

Nevertheless, the "generic" problem with $n = 2$ degrees of freedom is undoubtedly easier than the case of any $n \geq 3$. For, on the one hand, the isoenergetic reduction (§181) replaces the $2n$ -dimensional phase space, in the analytic case, by a $(2n - 1)$ -dimensional manifold for any n ; while, on the other hand, a theory of the possible compact 3-dimensional manifolds, though intricate enough in its de-

tails when no topologically admissible manifold is excluded, is to-day not so hopelessly remote as a corresponding theory for $n > 2$.

In this connection, mention may be made of a theorem of Poincaré which has no analogue at all in the higher-dimensional cases, and states that if there exists on a closed two-dimensional manifold a sheaf of curves which is free of singularities (in particular, of points of equilibrium), then the manifold must be topologically equivalent to a [non-orientable or orientable] torus. [For the tori occurring in §125, §196–§198, §215 (cf. also §121 bis, §127 bis), the non-orientable case is excluded by the fact that the systems considered reduce to the case of separated problems with a single degree of freedom, each of which determines a closed one-dimensional manifold (cf. (1₁), §185); so that the product space clearly is an orientable torus.]

In addition to the topological side of the issue, there arise several formal-analytical simplifications, if $n \geq 2$ is replaced by $n = 2$. In what follows, only these formal aspects of the case $n = 2$ can be considered. (As to a topological discussion, cf. the example of the projective space in §500).

§228. Let $n = 2$; so that (1), §155 may be written as

$$(1) \quad L = \frac{1}{2}(g_{11}x'^2 + 2g_{12}x'y' + g_{22}y'^2) + f_1x' + f_2y' + U,$$

where g_{ik} , f_i , U are six given functions of the coordinates $q_1 = x$, $q_2 = y$. One can assume without loss of generality that the $2 + 3$ functions $f_i(x, y)$, $g_{ik}(x, y)$ are expressible in terms of $1 + 1$ functions $f(x, y)$, $g(x, y)$ as follows:

$$(2_1) \quad f_1 = -yf, \quad f_2 = xf; \quad (2_2) \quad g_{11} = g, \quad g_{22} = g, \quad g_{12} = 0 \quad (g > 0).$$

First, one can replace f_1, f_2 by $f_1 + f_x, f_2 + f_y$, where $f = f(x, y)$ is arbitrary (§156). Hence, (2₁) requires merely that this $f(x, y)$ be chosen in a suitable way, namely so as to satisfy the linear partial differential equation $\omega = \frac{1}{2}\alpha$, where

$$(3) \quad 2\omega = xf_x + yf_y + 2f,$$

while α denotes $\partial f_2/\partial x - \partial f_1/\partial y$, a given function of (x, y) .

Next, it follows from (2₁), §155 and the assumption that the given functions $g_{ik}(x, y)$ are of class $C^{(2)}$, that $ds^2 = g_{11}dx^2 + 2g_{12}dxdy + g_{22}dy^2$ can be considered as the square of the line element on a surface which is embedded into a Euclidean 3-space and on which x, y are Gaussian parameters; and that this surface can be mapped upon a Euclidean plane (ξ, η) in such a way that the mapping is

locally topological and conformal. This means that if the "isothermic parameters" $\xi = \xi(x, y)$, $\eta = \eta(x, y)$ are used as Gaussian parameters on the surface, the invariant ds^2 appears as the product of a positive function $g = g(\xi, \eta)$ and of the Euclidean $d\sigma^2 = d\xi^2 + d\eta^2$. Hence, the admissibility of the normalization (2₂) follows from §95, if one writes x, y instead of ξ, η .

It should be mentioned that, in case of an isothermic parametrization (2₂) of the surface, the Theorema Egregium for the Gaussian curvature $K = K(x, y)$ on the surface is known to reduce to

$$(4) \quad K = \frac{1}{2}(g_x^2 + g_y^2 - gg_{xx} - gg_{yy})/g^3,$$

where $g = g(x, y) > 0$ is, of course, assumed to be of class $C^{(2)}$.

§229. Suppose* that $g(x, y) \equiv 1$. Then, from (1) and (2₁)–(2₂),

$$(5_1) \quad L = \frac{1}{2}(x'^2 + y'^2) + (xy' - yx')f + U;$$

$$(5_2) \quad X = x' - fy, \quad Y = y' + fx,$$

where X, Y denote the momenta, i.e., the partial derivatives of (5₁) with respect to x', y' . According to (5₁) and (3), the Lagrangian equations $[L]_x = 0$, $[L]_y = 0$ and the energy integral (3), §155 become

$$(6_1) \quad x'' - 2\omega y' = U_x, \quad y'' + 2\omega x' = U_y;$$

$$(6_2) \quad \frac{1}{2}(x'^2 + y'^2) - U(x, y) = h.$$

The Hamiltonian function belonging to (5₁) is, by §157,

$$(7_1) \quad H = \frac{1}{2}(X^2 + Y^2) - (xY - yX)f - V;$$

$$(7_2) \quad V = U - \frac{1}{2}(x^2 + y^2)f^2.$$

According to (5₁), (6₂), the function (2), §171 reduces to

$$(8) \quad M = (x'^2 + y'^2)^{\frac{1}{2}}(2U + 2h)^{\frac{1}{2}} + (xy' - yx').$$

§230. Introduce into (7₁) new coordinates ξ, η and momenta Ξ, H by means of the completely canonical transformation which is defined as the canonical extension of (the inverse of) a coordinate trans-

* This supposition involves, for a fixed value of the energy constant h , no loss of generality, as seen by identifying the function G of §180 with the function $g = g(x, y)$, and then applying the transformation of §180 to the Hamiltonian function $H = \frac{1}{2}(X^2 + Y^2)g^{-1} - \dots$ belonging to the Lagrangian function $L = \frac{1}{2}(x'^2 + y'^2)g + \dots$.

formation $x = x(\xi, \eta)$, $y = y(\xi, \eta)$. Suppose that this coordinate transformation is a conformal mapping of the (x, y) -plane upon the (ξ, η) -plane, i.e., that $z = x + iy$ is a regular analytic function $z = z(\zeta)$ of $\zeta = \xi + i\eta$, and $z_\zeta \neq 0$ in the domain under consideration; cf. §52. Then (7₁), §229 is transformed into (21), §52. Hence, if one applies the rule of §180 to $H = H(\Xi, H, \xi, \eta)$ by choosing $G = |z_\zeta(\xi + i\eta)|^2$, the Hamiltonian function (15), §180 belonging to a fixed value of the energy constant becomes

$$(9) \quad \bar{H} = \frac{1}{2}(\Xi^2 + H^2) - \frac{1}{2}(|z^2|_\xi H - |z^2|_\eta \Xi)f - \bar{V}(\xi, \eta; h),$$

where, according to (7₂) and §52,

$$(10_1) \quad \bar{V} = (U - \frac{1}{2}|z|^2 f^2 + h)|z_\zeta|^2; \quad (10_2) \quad 4|z_\zeta|^2 = |z^2|_{\xi\xi} + |z^2|_{\eta\eta}.$$

According to §180, the energy integral and the new time variable, \bar{t} , are

$$(11_1) \quad \bar{H} \equiv \bar{H}(\Xi, H, \xi, \eta; h) = \bar{h}, \quad \bar{h} = 0;$$

$$(11_2) \quad \bar{t} \equiv \bar{t}(t) = \int |z_\zeta(\zeta(t))|^{-2} dt.$$

If one denotes by dots differentiations with respect to \bar{t} and puts

$$(12_1) \quad \bar{U}(\xi, \eta; h) = |z_\zeta|^2(U + h); \quad (12_2) \quad \bar{\omega}(\xi, \eta) = |z_\zeta|^2 \omega,$$

the Lagrangian equations and their energy integral can be written as

$$(13_1) \quad \ddot{\xi} - 2\bar{\omega}\dot{\eta} = \bar{U}_\xi, \quad \ddot{\eta} + 2\bar{\omega}\dot{\xi} = \bar{U}_\eta;$$

$$(13_2) \quad \frac{1}{2}(\dot{\xi}^2 + \dot{\eta}^2) - \bar{U}(\xi, \eta; h) = 0.$$

In fact, the transformation rules of §157 show that the Lagrangian function $\bar{L} \equiv \bar{L}(\dot{\xi}, \dot{\eta}, \xi, \eta; h)$ belonging to (9) is

$$(14_1) \quad \bar{L} = \frac{1}{2}(\dot{\xi}^2 + \dot{\eta}^2) + \frac{1}{2}(|z^2|_\xi \dot{\eta} - |z^2|_\eta \dot{\xi})f + \bar{U};$$

$$(14_2) \quad \bar{U} = \bar{V} + \frac{1}{8}(|z^2|_\xi^2 + |z^2|_\eta^2)f^2.$$

Since $x + iy = z \equiv z(\zeta) = z(\xi + i\eta)$ satisfies the Cauchy-Riemann equations $x_\xi = y_\eta$, $x_\eta = -y_\xi$, the square sum () occurring in (14₂) readily reduces to $4|z|^2|z_\zeta|^2$; and so it is seen from (10₁) that the function (14₂) can be written in the form (12₁). Furthermore, it is easily found from (10₂) and from the definition $[L]_q = L_{q'} - L_q$, that the Lagrangian equations $[\bar{L}]_\xi = 0$, $[\bar{L}]_\eta = 0$ belonging to (14₁) reduce to (13₁), if $\bar{\omega}$ denotes the function $|z_\zeta|^2 f + \frac{1}{4}\{|z^2|_\xi f_\xi + |z^2|_\eta f_\eta\}$; so that $2\bar{\omega} = 2|z_\zeta|^2 f + |z_\zeta|^2(xf_x + yf_y)$ in view of the Cauchy-Rie-

mann equations. Hence, (12₂) follows from (3). Finally, (13₂) belongs to (13₁) in the same way as (6₂) does to (6₁), since the energy constant $\bar{h} = 0$, by (11₁).

§231. In view of (6₁), the product of the functions $\mp 2\omega(x, y)$ and of the velocity components x', y' is to be interpreted as a force of the Coriolis type. According to (3), the system is free of forces of this type if and only if the function $f(x, y)$ is homogeneous of degree -2 . Clearly, this will be the case if and only if $(xy' - yx')f(x, y)$ is the time derivative, G' , of a suitably chosen $G = G(x, y)$. According to §156, this will be the case if and only if the terms of (5₁) which are linear in x', y' can be omitted. Consequently, $\omega(x, y) \equiv 0$ if and only if $f(x, y) \equiv 0$. In other words, forces of the Coriolis type are absent if and only if the system is reversible (cf. §209–§210 and §155–§156).

In this case, (7₁)–(7₂) and (5₂) simplify to

$$(15_1) \quad H = \frac{1}{2}(X^2 + Y^2) - U(x, y); \quad (15_2) \quad X = x', \quad Y = y',$$

while (5₁), (6₁), (6₂) become

$$(16_1) \quad L = (x'^2 + y'^2) + U; \quad (16_2) \quad x'' = U_x, \quad y'' = U_y;$$

$$(16_3) \quad \frac{1}{2}(x'^2 + y'^2) - U = h,$$

(6₂) remaining unchanged (cf. §155). Similarly, from (8) and §172,

$$(17_1) \quad M = \{2(x'^2 + y'^2)(U(x, y) + h)\}^{\frac{1}{2}};$$

$$(17_2) \quad W = \int M dt, \quad (17_3) \quad W = \int (x'^2 + y'^2) dt.$$

Finally, §179 shows that, for a fixed value of the energy constant h , the integration problem of the Lagrangian equations (16₂) is equivalent to the problem of geodesics on the surface S_h on which the square $d\tilde{s}^2$ of the line element in terms of the Gaussian parameters x, y is

$$(18) \quad S_h: d\tilde{s}^2 = g^{(h)}(x, y)(dx^2 + dy^2), \quad \text{where } g^{(h)} = 2(U(x, y) + h).$$

Hence, if $K_h = K_h(x, y)$ denotes the Gaussian curvature on S_h , then

$$(19) \quad K_h = \frac{1}{4} \{ (U_x^2 + U_y^2) - (U + h)(U_{xx} + U_{yy}) \} / \{ U + h \}^3, \quad \text{by (4.)}$$

§231 bis. In (19), the denominator cannot vanish. In fact, it is understood (cf. §179) that the transition from (16₂)–(16₃) to (18) is valid only as long as $x'^2 + y'^2 \neq 0$, which means, by (16₃), that

$U(x, y) + h \neq 0$. Correspondingly, (18) shows that exactly those points of the Gaussian parameter plane (x, y) correspond to singular points of the surface S_h which lie on the manifold Z_h of zero velocity belonging to h . In fact, Z_h is, by §167, the set of points (x, y) at which $U(x, y) + h = 0$. Thus, if one excludes the trivial case $U(x, y) = \text{const.}$, then Z_h is a curve on the surface S_h or in the (x, y) -plane, it being understood that this curve, which need not be a connected set, may have a rather complicated structure and may contain isolated points or no point at all (cf. §168).

§232. Returning to the general case of §229, one sees that the integral (6₂) of (6₁) may be applied to a reduction of the system (6₁) of order four to a system of order three, if the energy constant h has a fixed value. Such a reduction can, for instance, be obtained by expressing y' from (6₂) as a function of $x, y, x'; h$ and then writing (6₁) as a system of three differential equations of the first order for x, y, x' ; a system in which h occurs as a fixed parameter. One can, however, carry out this isoenergetic reduction in a more symmetric manner, as follows:

For a given value of h , the energy equation (6₂) defines a "three-dimensional set" in the four-dimensional (x', y', x, y) -space. Let M_h denote that portion of this set at which $x'^2 + y'^2 \neq 0$; so that M_h is characterized by the conditions

$$(20) \quad M_h: \frac{1}{2}(x'^2 + y'^2) - U(x, y) - h = 0, \quad U(x, y) + h > 0, \\ (x'^2 + y'^2 \neq 0),$$

and is again a "three-dimensional set." It consists of those states (x', y', x, y) which satisfy the energy condition (6₂) for a fixed h but do not correspond to points of the set Z_h of zero velocity belonging to h . In other words, M_h consists of those points of the (x', y', x, y) -space which, when considered as representing initial conditions for (6₁), determine h as energy constant and a non-vanishing speed $(x'^2 + y'^2)^{\frac{1}{2}}$. The restriction imposed by the latter condition excludes only equilibrium points and cusps in the (x, y) -plane; cf. §169. In other words, the states (x', y', x, y) under consideration are those which belong to a given value of the energy constant h and are such that there exists in the (x, y) -plane a definite tangent with an inclination which is uniquely determined (mod 2π). Let w denote this inclination, so that $w = \arctan y'/x'$, where x' and y' do not vanish simultaneously; cf. (20). Accordingly, one can write

$$(21) \quad x' = v \cos w, \quad y' = v \sin w, \quad \text{where } v > 0; \quad w = \arctan y'/x'.$$

Now, it is clear that the set M_h defined by (20) can be parametrized in terms of the three independent variables x, y, v ; this parametrization being given by (21) if one puts

$$(22) \quad v = \{2(U(x, y) + h)\}^{\frac{1}{2}} > 0,$$

where $U + h > 0$, by (20). It is seen from (6₂) that (22) is the speed $(x'^2 + y'^2)^{\frac{1}{2}}$, expressed as a function of (x, y) for a fixed h .

The system of three differential equations alluded to can now be obtained as follows: Define three functions $\mathbf{x}, \mathbf{y}, \mathbf{w}$ of the three independent variables x, y, w and of the arbitrarily fixed parameter h , by placing

$$(23) \quad \begin{aligned} \mathbf{x} &= v \cos w, & \mathbf{y} &= v \sin w, \\ \mathbf{w} &= -2\omega + \{U_v \cos w - U_x \sin w\}/v, \end{aligned}$$

where v is the function (22) of x, y and h , while U_x, U_v and ω depend only on x, y ; cf. (3). Inasmuch as w' is, in view of (21), the ratio of $y''x' - x''y'$ and $x'^2 + y'^2$, substitution of x'', y'' and $x'^2 + y'^2$ from (6₁) and (6₂) shows that w' is, in virtue of (21), identical with the function \mathbf{w} defined by (23). Similarly, $x' = \mathbf{x}, y' = \mathbf{y}$ in view of (21), (22), (23). Thus,

$$(24) \quad x' = \mathbf{x}(x, y, w; h), \quad y' = \mathbf{y}(x, y, w; h), \quad w' = \mathbf{w}(x, y, w; h).$$

This means that those solutions of the system (6₁) of the fourth order which have the energy h are, in virtue of (21), identical with the solutions of the system (24) of the third order, if one excludes states $x'(t)^2 + y'(t)^2 = 0$ of zero velocity and defines the functions $\mathbf{x}, \mathbf{y}, \mathbf{w}$ by (23), (22).

It is easily verified from (22) by partial differentiations of the three functions (23) with respect to x, y, w , respectively, that

$$(25) \quad \mathbf{x}_x(x, y, w; h) + \mathbf{y}_y(x, y, w; h) + \mathbf{w}_w(x, y, w; h) \equiv 0.$$

§232 bis. If $k = k(t)$ and $s = s(t)$ denote the curvature and the arc length along the positively oriented* solution path $x = x(t)$,

* The path will be considered as "positively oriented" if the arc length $s = s(t)$ increases with t ; while the curvature k is defined, with reference to increasing s , as dw/ds , the angle w in (21) being the inclination of the tangent towards the positively oriented x -axis. Notice that, in contrast with the Gaussian curvature of a surface, the curvature of a curve is defined in an in-

$y = y(t)$ of energy h in the (x, y) -plane, the system (25) of three equations of the first order, which is valid if $x'^2 + y'^2 \neq 0$, is equivalent to the pair of "intrinsic" equations

$$(26) \quad k = (-2\omega v + U_y \cos w - U_x \sin w)/v^2; \quad s' = v,$$

(the first of which is, in view of the definition of curvature, a differential equation of the second order). In fact, since (22), (6₂) show that v is the speed $(x'^2 + y'^2)^{\frac{1}{2}} > 0$, it is clear from $ds = (dx^2 + dy^2)^{\frac{1}{2}}$ that $s' = v$. On the other hand, the definitions of k and of the angle w in (21) imply that $k = dw/ds$, i.e., $k = w'/s' \equiv w'/v$; whence the representation (26) of k follows by substituting w' from (24), and then w from (23).

§233. As an application, consider a solution path which is periodic. Let h denote the energy and τ the period of this path, which represents a closed curve C in the (x, y) -plane. Suppose for simplicity that C has no self-intersections ("loops"), i.e., that C is a Jordan curve, and let D denote the bounded Jordan domain bordered by C . Suppose further that the given function (3) which occurs in the Lagrangian equations (6₁) is $\omega(x, y) \equiv 1$. Suppose finally that neither C nor its interior D contains a point (x, y) of the set of zero velocity belonging to h , i.e., that the inequality (22) is satisfied on and within C ; so that $\Delta^2 \log v$, where $\Delta^2 F$ denotes the Laplacian $F_{xx} + F_{yy}$, is a continuous function of (x, y) on $C + D$.

It will be shown that, if these assumptions are satisfied, the period τ of the given solution can be expressed in terms of the double integral

$$(27) \quad I = \iint_D \Delta^2 \log v(x, y; h) dx dy.$$

Consider first the case in which C surrounds the origin of the (x, y) -plane, and the orientation of C is counter-clockwise; so that the angular coordinate $w = w(t)$ defined by (21) increases by 2π during a t -interval of length $\tau = \text{period} > 0$. Thus,

$$(28) \quad 2\pi = \int_0^\tau w' dt = -2\tau \cdot 1 + \int_0^\tau \{U_y \cos w - U_x \sin w\} v^{-1} dt,$$

variant manner only in absolute value. Correspondingly,

$$k = (y''x' - x''y') : (x'^2 + y'^2)^{\frac{3}{2}}$$

contains a square root (≥ 0), while $k = dw/ds$ changes its sign if one allows s to decrease, instead of to increase, with t .

by (24), (23), where $\omega \equiv 1$ by assumption.* Hence, from $\mathbf{v} = ds/dt$,

$$(29) \quad -2\pi - 2\tau = \int_C U_n \mathbf{v}^{-2} ds,$$

since $-\{U_y \cos w - U_x \sin w\}$ is, according to (21), the (exterior) normal derivative of $U = U(x, y)$ along the positively oriented path C . But $\mathbf{v}_n = U_n \mathbf{v}^{-1}$, by (22). Hence, the integrand of (29) can be written as $\mathbf{v} \mathbf{v}_n \mathbf{v}^{-2} \equiv (\log \mathbf{v})_n$; so that the line integral (29) is, by Green's theorem, identical with the double integral (27). Consequently, (29) can be written as

$$(30) \quad \tau = -\frac{1}{2}I - \pi.$$

This is the desired representation of the period τ . It is clear from the proof how (29) must be modified when the simple closed curve C , instead of being direct, is retrograde about the origin, or when it does not surround the origin.

Thus far it has been assumed that (22) is satisfied not only on C but also in the interior, D , of C . If D contains a finite number of points $(x, y) = (a, b)$ at which the function (22) of (x, y) vanishes, for the fixed value of h , in the order of r , where $r^2 = (x - a)^2 + (y - b)^2$, then, before applying Green's theorem, one has to exclude in (27) circles of radius ϵ about these points, and then let $\epsilon \rightarrow 0$. This leads, in a manner well known from the elements of the theory of the logarithmical potential, to a modification of the relation which connects τ and I , the modification consisting in an additive multiple of π . Obviously, the same holds if r is replaced by r^m , where $m > 0$, and also if the function (22), i.e., the force function $U(x, y)$, becomes infinite at $(x, y) = (a, b)$ in some order m . The latter case occurs in the restricted problem of three bodies.

§233 bis. It is clear from the above proofs that if C , instead of being a simple closed curve, consists of a finite system of loops, then the representation of τ in terms of I must be modified by an index number, determined by the ramifications of C . The resulting existence problems then involve the relations of Birkhoff concerning the critical points of functions of two variables; relations which were de-

* In the reversible case, where $\omega \equiv 0$, one has $-2\tau\omega = 0$ for every τ ; so that τ does not occur in (29) and, instead of a connection between the period τ and the integral (27), one obtains the simplest case of the Gauss-Bonnet theorem; cf. the end of §231.

veloped by him, and subsequently generalized by Morse for the multi-dimensional case, precisely in this connection.

§234. Let $(x(t), y(t))$ be a solution of (6₁), with energy h :

$$(31_1) \quad x = x(t), \quad y = y(t); \quad (31_2) \quad x'^2 + y'^2 = 2(U(x, y) + h),$$

and suppose, for simplicity, that the "angular velocity" $\omega = \omega(x, y)$ defined by (3) is a constant; so that $f = \omega$, and so, by (1) and (2₁)–(2₂),

$$(32_1) \quad x'' - 2\omega y' = U_x, \quad y'' + 2\omega x' = U_y;$$

$$(32_2) \quad L = \frac{1}{2}(x'^2 + y'^2) + (xy' - yx')\omega + U.$$

The Jacobi equations which define the displacements $x = x(t)$, $y = y(t)$ of (31₁) are represented by

$$(33) \quad x'' - 2\omega y' = U_{xx}x + U_{xy}y, \quad y'' + 2\omega x' = U_{xy}x + U_{yy}y \\ (U_{xx} = U_{xx}(x(t), y(t)), \dots).$$

In fact, it is seen from (32₂) that the Lagrangian function L defined by (21₂), §101 is

$$(34_1) \quad L = \frac{1}{2}(x'^2 + y'^2) + (xy' - yx')\omega + U(x, y; t);$$

$$(34_2) \quad U = \frac{1}{2}(U_{xx}x^2 + 2U_{xy}xy + U_{yy}y^2),$$

where U_{xx}, \dots denote the known functions $U_{xx}(t) = U_{xx}(x(t), y(t))$, \dots of t along the given solution (31₁) of (32₁), and x, y are the components of the vector κ occurring in (21₂), §101. Now, it is clear from (34₁)–(34₂) that the Jacobi equations $[L]_x = 0$, $[L]_y = 0$ of §101 take the form (33).

The Jacobi equations (33) have the linear integral

$$(35_1) \quad x'(t)x' + y'(t)y' - U_x(t)x - U_y(t)y = h; \quad (35_2) \quad h = \text{const.}$$

In fact, (35₁) is, in view of (21₁)–(21₂), §101, identical with (22), §101.

According to (35₁) and the definition in §102, the isoenergetic displacements of (31₁) are those solutions $(x(t), y(t))$ of (33) for which one has, as an identity in t ,

$$(36) \quad x'x' + y'y' = U_x x + U_y y, \quad (\text{i.e., } h = 0).$$

By the end of §102, a particular isoenergetic displacement of (31₁)

$$(37) \quad x = x'(t), \quad y = y'(t), \quad (h = 0).$$

§235. Suppose that, on the t -interval under consideration, the given solution (31_1) of (32_1) does not reach its manifold of zero velocity, i.e., that (31_2) does not vanish on this t -interval. This assumption is identical with that of §232–§233 and excludes the case where (31_1) is, in the (x, y) -plane, either a single point or a curve which has a cusp for some t . Hence, one can define, with reference to any given solution $x = x(t)$, $y = y(t)$ of (33) , a function $n = n(t)$ by placing

$$(38_1) \quad n = d/v; \quad (38_2) \quad d = x'y - y'x; \quad (38_3) \quad v = (x'^2 + y'^2)^{\frac{1}{2}} > 0.$$

According to the definitions (38_2) , (38_3) of the functions d , v of t , the function (38_1) is, for every t , the projection of the displacement $(x(t), y(t))$ on the normal of the given solution path (31_1) of (32_1) , this normal being oriented by the choice of the sign of the square root (38_3) . Correspondingly, a function $n = n(t)$ is called a normal displacement of (31_1) if (33) possesses a solution $(x(t), y(t))$ by means of which $n(t)$ is representable in the form (38_1) – (38_2) . If, in particular, $n(t)$ belongs to a displacement $(x(t), y(t))$ for which the integration constant (35_2) vanishes, then $n(t)$ is called an isoenergetic normal displacement of (31_1) .

§236. It will be shown that one can calculate, for any given solution path (31_1) which satisfies the assumption (38_3) , a unique continuous scalar function $\kappa = \kappa(t)$ in such a way that a scalar function $n = n(t)$ is an isoenergetic normal displacement of (31_1) if and only if it satisfies the linear differential equation

$$(39) \quad n'' + \kappa(t)n = 0; \quad \text{cf. (45) and (44) below.}$$

This seems to be a paradox, since the general solution of (39) , where $\kappa(t)$ is given, contains two arbitrary integration constants, while the isoenergetic displacement $(x(t), y(t))$ which defines $n(t)$ depends on three such constants (the general solution of (33) depends on four integration constants one of which is fixed by the isoenergetic assumption that (35_2) vanishes). The explanation is that, according to (38_1) – (38_2) , the trivial solution $n(t) \equiv 0$ of (39) belongs not only to the trivial solution $x(t) \equiv 0$, $y(t) \equiv 0$ of (33) and (36) but also to the isoenergetic displacement which results if one multiplies both functions (37) by an arbitrary constant factor c ; so that this c is the missing integration constant. In fact, (37) is not the trivial solution $x(t) \equiv 0$, $y(t) \equiv 0$, since otherwise (31_1) were an equilibrium solution, and this is excluded by (38_3) .

In the construction of the coefficient function, $\kappa(t)$, of (39) by means of appropriate differentiations and eliminations, use will be made of the relation

$$(40) \quad \left(\frac{1}{2}v^2\right)'d' - [x'y' - y'x']v^2 = \{x'x' + y'y' - x''x - y''y\}(x'y'' - y'x'') + (x''^2 + y''^2)d.$$

This is an algebraic identity, since, from (38₂)–(38₃),

$$(41_1) \quad d' = x''y - y''x + x'y' - y'x'; \quad (41_2) \quad \left(\frac{1}{2}v^2\right)' = x'x'' + y'y''.$$

§237. Suppose that $(x(t), y(t))$ is an isoenergetic displacement.

On differentiating (41₁) with respect to t and then expressing x'' , y'' and x''' , y''' in d'' by using (33) and the relations which result by differentiation of (32₁), one readily obtains

$$(42) \quad d'' = -2\omega\{x'x' + y'y' - x''x - y''y\} + (U_{xx} + U_{yy})d + 2[x''y' - y''x'],$$

since $\omega = \text{const.}$ and $d = x'y - y'x$. Substituting $[x''y' - y''x']$ from (42) into (40), and noting that the expression $\{ \}$ which occurs in both (42) and (40) is, in view of (36) and (32₁), simply $2\omega(x'y - y'x) \equiv 2\omega d$, one sees that

$$(43) \quad v^4(d'/v^2)' + ud = 0,$$

where $u = 2(x''^2 + y''^2) - 4(x''y' - y''x')\omega + (4\omega^2 - U_{xx} - U_{yy})v^2$. Thus, on expressing x'' , y'' in u by means of (32₁), one obtains

$$(44) \quad u = 2(U_x^2 + U_y^2) + 4(U_x y' - U_y x')\omega + (4\omega^2 - U_{xx} - U_{yy})v^2, \\ \text{where } v^2 = 2(U + h),$$

by (38₃) and (31₂). Finally, on substituting $d = vn$ from (38₁) into (43), one sees that the homogeneous linear differential equation (43) for d appears in the self-adjoint form (39); the resulting explicit representation of $\kappa = \kappa(t)$ in terms of the functions (44), (38₃) of t being

$$(45) \quad \kappa = (v''v - 2v'^2 + u)/v^2.$$

§237 bis. It remains to prove the converse, namely, that there exists for every given solution $n(t)$ of (39) an isoenergetic displacement $(x(t), y(t))$ which satisfies (38₁)–(38₂); or, what is the same thing, that there exists for every given solution $d(t)$ of (43) a pair of functions $x(t), y(t)$ which satisfy (33), (36) and (38₂).

To this end, let d^0, x'^0, U_x^0, \dots denote the values which the functions $d(t), x'(t), U_x(x(t), y(t)), \dots$ of t attain at some fixed $t = t^0$, where (31_1) and $d(t)$ are given solutions of (32_1) and (43) , respectively. Since (38_3) implies that at least one of the two numbers x'^0, y'^0 does not vanish, one can assume that $x'^0 \neq 0$. Starting with the given numbers d^0, d'^0, x'^0, \dots , determine four numbers x^0, y^0, x'^0, y'^0 which satisfy the three linear conditions

$$(46_1) \quad x'^0 y^0 - y'^0 x^0 = d^0;$$

$$(46_2) \quad x''^0 y^0 - y''^0 x^0 + x'^0 y'^0 - y'^0 x'^0 = d'^0;$$

$$(46_3) \quad x'^0 x'^0 + y'^0 y'^0 = U_x^0 x^0 + U_y^0 y^0.$$

This is possible, since, on choosing x^0 arbitrarily, one sees that (46_1) – (46_3) becomes a system of three linear equations for y^0, x'^0, y'^0 which has the determinant $-(x'^0 x'^0 + y'^0 y'^0) x'^0 \neq 0$.

Let $x = x(t), y = y(t)$ be that solution of (33) which belongs to the initial values x^0, y^0, x'^0, y'^0 , assigned to some $t = t^0$. Then (36) is satisfied, since (35_1) is an integral of (33) and the constant (35_2) vanishes, by (46_3) . Hence, on placing $\bar{d}(t) = x'y - y'x$, one can conclude from §237 that $d = \bar{d}(t)$ is a solution of (43) . Furthermore, $\bar{d}^0 = d^0, \bar{d}'^0 = d'^0$, by (46_1) – (46_2) . Since the differential equation (43) of the second order has only one solution which attains given initial values d^0, d'^0 , it follows that $d(t) \equiv \bar{d}(t)$; so that the given solution $d(t)$ of (43) is representable in the desired form (38_2) .

§238. Leaving aside the investigation of the displacements $(x(t), y(t))$ of (31_1) , suppose that the speed $v(t) = (x'^2 + y'^2)^{\frac{1}{2}}$ of the given solution (31_1) of (32_1) vanishes at some t , say at $t = 0$, without vanishing for every t ; so that the path (31_1) has at $t = 0$ a cusp. Denoting by ζ^0 the initial value $\zeta(0)$ of any function $\zeta(t)$ of t , one sees from §168 that the point (x^0, y^0) of the (x, y) -plane is on the curve $U(x, y) = -h$ of zero velocity, where $h = -U(x^0, y^0)$. Furthermore, §166 shows that this curve through (x^0, y^0) has at (x^0, y^0) a definite normal; while §170 states that this curve reflects the path in the transversal direction. In the present case, the transversal to $U(x, y) = -h$ at (x^0, y^0) is the normal, since the g_{ik} of (32_2) are Euclidean. Accordingly, the path (31_1) becomes, as $t \rightarrow \pm 0$, tangent to the normal of the curve $U(x, y) = -h$ through (x^0, y^0) and lies, for small $t \leq 0$, on one and the same side of this curve. Let the positive normal of $U(x, y) = -h$ at (x^0, y^0) be defined as that half of the tangent of the cusp of (31_1) which lies on the same side of the

curve $U(x, y) = -h$ as the cusp; this half of the tangent to the cusp of (31₁) will also be called positive.

Suppose that $\omega(x, y) \equiv 1$; so that (32₁) can be written as

$$(47) \quad x'' = 2y' + U_x, \quad y'' = -2x' + U_y.$$

It will be shown that, if an observer moves along the path (31₁) in the direction which the path exhibits when t increases, then he will see the positive tangent of the cusp on his left both before ($t < 0$) and after ($t > 0$) passing through the point (x^0, y^0) . This implies, in particular, that the positive tangent is an inner tangent of the cusp.

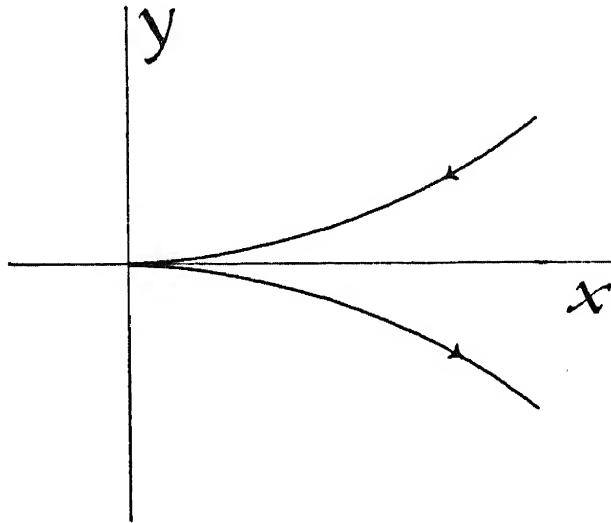


FIG. 2

Since (47) clearly remains unchanged under both a constant translation and a constant rotation of the (x, y) -plane, one can assume that the point (x^0, y^0) is the origin $(0, 0)$, and that the positive normal of the curve $U(x, y) = -h$ through $(x, y) = (0, 0)$ is the positive half of the x -axis. Then $U_y^0 = 0$; hence, $U_x^0 \neq 0$, since the simultaneous vanishing of U_x^0 and U_y^0 is, by §165, possible only in case of an equilibrium solution. Since x' and y' vanish at the cusp $t = 0$, one sees from (47) that $x''^0 = U_x^0 \neq 0$, $y''^0 = 0$. But the positive tangent of the cusp is the positive half of the x -axis; so that $x(t) > 0 (= x^0)$ for small $t \leq 0$, and so Taylor's formula shows that the non-vanishing constant x''^0 is positive. Thus,

$$(48_1) \quad x^0 = 0, \quad y^0 = 0; \quad (48_2) \quad x'^0 = 0, \quad y'^0 = 0;$$

$$(48_3) \quad x''^0 = U_x^0 > 0, \quad y''^0 = U_y^0 = 0.$$

On differentiating the second equation (47) with respect to t at

$t = 0$, one sees from (48₂)–(48₃) that $y''''^0 = -2x''^0$. Hence, from (48₁)–(48₃) and by Taylor's formula,

$$(49) \quad \begin{aligned} x(t) &= \alpha t^2 + o(t^2), \\ y(t) &= -\frac{2}{3}\alpha t^3 + o(|t|^3), \quad \text{where } \alpha = \text{const.} \gtrless 0, \end{aligned}$$

and $t \rightarrow \pm 0$. Clearly, (49) implies the orientation rule which was to be proved, and shows also that, to the first approximation, the two branches of the cusp are identical semi-cubical parabolas.

§239. According to §85, the knowledge of the displacements $(x(t), y(t))$ of (31₁) leads to an approximate determination of those solutions of (32₁) which belong to initial values close to those of (31₁). This is the practical significance of the result of §236. For, on the one hand, §236 reduces to (39) the determination of a family of solutions of (33) which depend on three integration constants; and, on the other hand, the general solution of the system (33) of the fourth order, i.e., the introduction of a fourth integration constant, requires merely a quadrature (this fourth integration constant is, of course, the deviation, (35₂), from an isoenergetic displacement).

However, §236 breaks down in the neighborhood of any fixed t , say $t = 0$, which is such that the speed $v(t) = (x'^2 + y'^2)^{\frac{1}{2}}$ of (31₁) vanishes at $t = 0$. If $v(t)$ vanishes for every t , there arises no difficulty (although (38₁), (39), (43)–(45) become meaningless for every t). In fact, if (31₁) is an equilibrium solution, the determination of its displacements $x(t), y(t)$ is, according to §89, a trivial task.

There remains to be considered the case where $v(t)$ vanishes at $t = 0$ but not at every t . Then the differential equation (43) and, correspondingly, the coefficient (45) of (39), acquires a singularity at $t = 0$ (which agrees with the geometrical meaning of (38₁)–(38₃), since the path (31₁) has a cusp at $t = 0$). Hence, in order to discuss the approximate behavior of solution paths which lie close to the given solution path (31₁) having at $t = 0$ a cusp, a direct procedure is necessary.

§240. To this end, let the given solution path (31₁) be the same as in §238. In order to obtain solution paths of (47) which belong to a slightly different set of initial values, replace the initial conditions (48₁)–(48₂) by

$$(50_1) \quad x^0 = 0, \quad y^0 = 0; \qquad (50_2) \quad x'^0 = 0, \quad y'^0 \gtrless 0,$$

where y'^0 is an arbitrary small integration constant, it being under-

stood that $y'^0 = 0$ belongs to the cuspidal solution of Fig. 2. It is seen by an obvious repetition of the calculations which led from (48₁)–(48₂) to (49), that in case of the integration constants (50₁)–(50₂) one has, as $t \rightarrow \pm 0$,

$$(51) \quad \begin{aligned} x(t) &= (y'^0 + \alpha)t^2 + y'^0\beta t^3 + o(|t|^3), \\ y(t) &= y'^0 t + (y'^0\gamma - \frac{2}{3}\alpha)t^3 + o(|t|^3), \end{aligned}$$

where the numbers α, β, γ depend only on the numbers $U_x^0, U_{xy}^0, U_{yy}^0$, and are, therefore, independent of the parameter y'^0 . In particular, $\alpha (= \frac{1}{2}U_x^0)$ is the same number in (51) as in the particular case $y'^0 = 0$; so that $\alpha > 0$, by (49).

Suppose, for instance, that y'^0 is chosen as a small positive number. Then, since $\alpha > 0$, it is clear from (51) that $y(t)$ vanishes not only at $t = 0$ but also at a small positive and at a small negative value of t ; values which are, for small y'^0 , approximately given by the roots of the quadratic equation $y'^0 + (y'^0\gamma - \frac{2}{3}\alpha)t^2 = 0$ and lie, therefore, very close to

$$(52) \quad \pm \left\{ -y'^0 / (y'^0\gamma - \frac{2}{3}\alpha) \right\}^{\frac{1}{2}} \sim \pm \left\{ \frac{3}{2}y'^0 / \alpha \right\}^{\frac{1}{2}}, \quad \text{as } y'^0 \rightarrow +0; \\ (\alpha = \text{const.} > 0).$$

Since $y(t)$ vanishes at $t = 0$ and at two additional t close to the small values (52), it is clear that, for small values of the integration con-

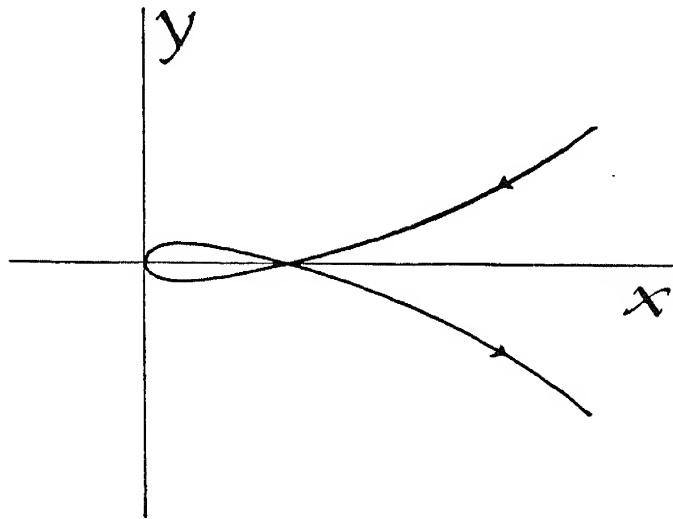


FIG. 3

stant $y'^0 > 0$, the solution (51) of (47) determined by (50₁)–(50₂) has a small loop which disappears in the cusp of Fig. 2, as $y'^0 \rightarrow +0$.

The direction in which the loop is described is clear from (51) and, for reasons of continuity, from the rule of §238 also.

The situation becomes clearer by observing that the curve of zero velocity belonging to the energy constant h of (51) is not the same for $y'^0 = 0$ as for small $y'^0 > 0$. In fact, substitution of (51) into (31₂) shows that $h = h(y'^0)$ is a continuous function which attains for the cuspidal case $y'^0 = 0$ an extremal value. This holds also when the small integration constant y'^0 is allowed to be negative; in which case the solution (51) of (47) has neither a loop nor a cusp close to $(x, y; t) = (0, 0; 0)$. Since the curve of zero velocity is changing with y'^0 , it becomes understandable why (51) has a cusp only when $y'^0 = 0$.

CHAPTER IV

THE PROBLEM OF TWO BODIES

The solution paths	§241–§257
The anomalies	§258–§273
Expansions of the elliptic motion into Fourier series	§274–§284
Expansions according to powers of the eccentricity	§285–§299
Synodical coordinates	§300–§312

The Solution Paths

§241. Consider the case $U(r) = r^{-1}$ of §218. Thus,

$$(1) \quad \begin{aligned} L &= \frac{1}{2}(x'^2 + y'^2) + r^{-1}, \quad \text{where } r = (x^2 + y^2)^{\frac{1}{2}}; \text{ hence,} \\ H &= \frac{1}{2}(x'^2 + y'^2) - r^{-1}, \end{aligned}$$

the momenta $L_{x'}$, $L_{y'}$ reducing to the velocities x' , y' . According to (11₁)–(11₃), §211, the equations of motion and the conservation of energy and of angular momentum are

$$(2_1) \quad x'' = -xr^{-3}, \quad y'' = -yr^{-3};$$

$$(2_2) \quad \frac{1}{2}(x'^2 + y'^2) - r^{-1} = h; \quad (2_3) \quad xy' - yx' = c.$$

The discussion of the curve representing an arbitrary solution path of (2₁) in the configuration plane (x, y) may be carried out as follows:

First, (2₁), (2₃), where $r = (x^2 + y^2)^{\frac{1}{2}}$, imply that $cy'' = (xr^{-1})'$, $cx'' = (-yr^{-1})'$. Thus, if A, B denote integration constants, then $cy' = xr^{-1} + A$, $cx' = -yr^{-1} - B$. Hence, from (2₂)–(2₃),

$$(3_1) \quad A^2 + B^2 = 1 + 2hc^2; \quad (3_2) \quad c^2 = Ax + By + r; \quad r = (x^2 + y^2)^{\frac{1}{2}}.$$

Clearly, (3₂) is the equation of (a branch of) a conic.

In order to simplify the discussion of this conic, replace the integration constants (h, c) of energy and of angular momentum by equivalent integration constants (a, e) in the case $c^2 \geq 0$, $h \neq 0$, and $c^2 \geq 0$ by an integration constant p in the remaining case $h = 0$, by placing

$$(4) \quad \begin{aligned} e &= (1 + 2hc^2)^{\frac{1}{2}} \geq 0, \\ a &= (-2h)^{-1}, \text{ if } h \neq 0; \text{ and } p = c^2 \geq 0, \text{ if } h = 0, \end{aligned}$$

where the radicand of e cannot be negative.* In fact, (3₁) shows

* It is easy to verify that this limitation of the integration constants h, c

that only those values of the angular momentum constant c are compatible with a given value of the energy constant h for which $1 + 2hc^2 \geq 0$.

§242. It is clear from (3₁) that the branch of a conic (3₂), which can degenerate into a half-line or a segment through $(x, y) = (0, 0)$, has at $(x, y) = (0, 0)$ a focus, and that the conic does not degenerate if and only if $c \neq 0$. (That the rectilinear motion is, for $h \leq 0$, characterized by $c = 0$, is clear from (2₃) also.) Furthermore, it is easily verified from (3₁) and (4) that $2a \geq 0$ and $e \geq 0$ are the major axis and eccentricity or $p \geq 0$ is the parameter of the conic (3₂) according as the other focus is not or is at infinity. Since a and $-h^{-1}$ are, by (4), of the same sign, it follows that, no matter what is the value of c , the elliptic, hyperbolic and parabolic cases are characterized by $h < 0$, $h > 0$ and $h = 0$, respectively. Hence, the path is closed in the (x, y) -plane if $h < 0$; in which case the period of $x = x(t)$, $y = y(t)$ is, by §160 bis, proportional to $|h|^{-\frac{3}{2}}$, since $\beta^{-1} - 1 = -\frac{3}{2}$ in the present case $U = (x^2 + y^2)^{-\frac{1}{2}}$. It is clear from (4) that the ellipse becomes a circle if and only if

$$(5) \quad 1 + 2hc^2 = 0 \quad [\text{cf. (18}_1\text{), §216}].$$

The general connection (4) between the integration constants $h \leq 0$, $c \geq 0$ and the geometrical data $a \geq 0$, $e \geq 0$ or $p \geq 0$ is seen to be such that

$$(6) \quad e > 0 \text{ unless } (-2h)^{-1} = a = c^2 > 0 = e;$$

$$(7) \quad \text{if } c \neq 0, \text{ then } e \leq 1, a \geq 0 \text{ for } h \leq 0, \text{ while } p > 0 \text{ for } h = 0;$$

$$(8) \quad \text{if } c = 0, \text{ then } e = 1 \text{ for } h \leq 0, \text{ while } p = 0 \text{ for } h = 0.$$

That only the square of c can occur in (4), is clear from the fact that, on changing c to $-c$, one changes merely the orientation of the motion, but not the path; cf. §214.

It is easily verified* that if $h > 0$ and $c \neq 0$, the path is that of the two branches of the hyperbola which shows its concavity towards the focus $(x, y) = (0, 0)$.

§243. According to (2₂), the equation of the curve of zero velocity

is equivalent to the limitation imposed on a path of energy h by the manifold of zero velocity belonging to h ; cf. §243.

* In fact, the numerator, $U_y \cos w - U_x \sin w$ ($\omega \equiv 0$), of the curvature (26), §232 bis reduces, in case of an arbitrary $U = U(r)$ in (11₁), §211, to $(y \cos w - x \sin w)U_r/r$, where $w = w(t)$ is the inclination of the path; and $U_r < 0$ in case of attraction.

belonging to a given h is $(x^2 + y^2)^{-\frac{1}{2}} = -h$; so that this curve exists only when $h < 0$, in which case it is the circle of radius $-h^{-1}$ about $(x, y) = (0, 0)$. This radius is $2a$, where a is the radius of a concentric circle which represents the circular path of energy h . In fact, the major axis, $2a$, of an elliptic path is $-h^{-1}$, by (4). Since $2a$ is independent of c , it is also seen that a path of energy $h < 0$ has a point on its circle of zero velocity only when the eccentricity $e = 1$, i.e., when the ellipse degenerates into a segment represented by a radius of the curve of zero velocity. If, on the other hand, $e < 1$, the ellipse is in the interior of its curve of zero velocity, since then the focus $(x, y) = (0, 0)$ is in the interior of the ellipse.

All this agrees with what was proved about cusps in §169–§170. Notice, however, that the general theory is not applicable to solution arcs which reach the focus $(x, y) = (0, 0)$, since (2₁) then has a singularity.

§244. If $h \leq 0$ is arbitrarily fixed, the square of the line element on the surface S_h of §212 is the product of g and $dx^2 + dy^2 \equiv dr^2 + r^2 d\phi^2$, where $g = 2(r^{-1} + h)$. Hence, the singularities of the surface S_h of revolution are the parallel circles (or points) along which either $g = \infty$ or $g = 0$, i.e., either $(x, y) = (0, 0)$ or $r^{-1} = -h$. The singularities of the first kind on S_h occur for any $h \leq 0$, while singularities of the second kind, which represent the curve of zero velocity, only when $h < 0$.

Barring the singularities of S_h , one sees from (13), §212 that the Gaussian curvature is $-\frac{1}{4}h(1 + rh)^{-3}$, since $U = r^{-1}$. But (2₂), §241 implies that $1 + rh \equiv r(r^{-1} + h)$ is positive. Hence, the Gaussian curvature on S_h is everywhere positive, zero or negative according as $-\frac{1}{4}h \geq 0$. In other words, every non-singular point of S_h is elliptic, parabolic or hyperbolic in the sense of differential geometry according as the energy constant h belongs, in the sense of (4), to elliptic, parabolic or hyperbolic paths in the (x, y) -plane (in particular, the metric of S_h is Euclidean if and only if $h = 0$). It follows that if $h \geq 0$, there cannot exist conjugate points on the geodesics of S_h .

§245. If $s = s(t)$ denotes the arc length along a solution path $x = x(t)$, $y = y(t)$ of fixed energy $h (\leq 0)$ in the (x, y) -plane, then $s'^2 = x'^2 + y'^2$; so that, from (17₃), §231,

$$(9) \quad W' = s'^2 \equiv x'^2 + y'^2. \qquad 10) \quad W = \int_{P^0}^P s'^2 d\bar{t},$$

where W has the same meaning as in §99, with the understanding that the integration in (10) is extended along the given solution path between a fixed and a variable point, $P^0: (x(t^0), y(t^0))$ and $P: (x(t), y(t))$.

It will be shown that, barring the case $c = 0$ of a rectilinear path, one has for the function (10) of t a simple geometrical interpretation in all three cases $h \not\leq 0$. If $h \leq 0$, i.e., if the conic has two foci O, F (where O is the origin of the (x, y) -plane and coincides with F in the circular case), the interpretation in question is analogous to the interpretation of the constant (2₃) as the two-fold areal velocity about the focus O , where the parabolic case $h = 0$ is not excluded.

For $h \not\leq 0$, let $\sigma = \sigma(t)$ denote the area of the sector bordered by the arc (P^0, P) of the path $x = x(t), y = y(t)$ and the radii vectores which connect the points $P^0 = P(t^0)$ and $P = P(t)$ with the focus O ; so that σ' is the areal velocity about the origin, and so $2\sigma' = c$, where $c \neq 0$ by assumption. Furthermore, if $l = l(t)$ denotes the length of the perpendicular drawn from O to the line which touches the path at $P = P(t)$, then $d\sigma = lds$, since $s = s(t)$ is the arc length; so that $\sigma' = ls'$.

Exclude, for a moment, the case $h = 0$ of a parabola. Let $\bar{\sigma} = \bar{\sigma}(t)$ and $\bar{l} = \bar{l}(t)$ denote the functions which result if, without changing P^0, P and s , one replaces the focus O by the focus F in the definitions of $\sigma = \sigma(t)$ and $l = l(t)$. Then, corresponding to $\sigma' = ls'$, one has $\bar{\sigma}' = \bar{l}s'$, and so $\sigma'\bar{\sigma}' = \bar{l}ls'^2$. But the product $\bar{l}l$ is a non-vanishing constant by a property of ellipses and hyperbolas; so that $\sigma'\bar{\sigma}'$ is proportional to s'^2 , and so, by (9), to W' . Since $2\sigma'$ was seen to be a non-vanishing constant, it follows that the functions $\bar{\sigma}'$ and W' of t are proportional.

Accordingly, while the area $\sigma = \sigma(t)$ referred to the focus O is, for $h \not\leq 0$, proportional to t , the area $\bar{\sigma} = \bar{\sigma}(t)$ referred to the focus F is proportional to the function W of t , if $h \leq 0$. It is easily verified that this interpretation of $W = W(t)$ holds in the limiting case $h = 0$ of a parabola also; in which case $\bar{\sigma} = \bar{\sigma}(t)$ has to be defined as the area bordered by the arc (P^0, P) of the parabola, the axis of the parabola, and the perpendiculars drawn from P^0 and P to the axis of the parabola.

§246. In what follows, the origin $(x, y) = (0, 0)$ will be referred to as the focus O also in the rectilinear case $c = 0$. The other focus, F , which exists only in the non-parabolic cases $h \leq 0$, will be called (also in the circular case $O = F$) the "empty" focus; O being thought of as containing the attracting mass.

If $P^0 = P(t^0)$ and $P = P(t)$ denote the points of the solution path which belong to a fixed t^0 and a variable t , one and the same pair P^0, P can belong to two different pairs t^0, t . Let this ambiguity (which will arise only when either $h < 0$ or $c = 0$) be eliminated by the requirement that t is the first date which follows t^0 and belongs to the position P .

§247. Introducing polar coordinates $q_1 = r, q_2 = \phi$ and applying, for instance, the last of the rules (22) of §116 bis, one sees that the time, $t - t^0$, which elapses between the two positions P^0, P depends only on the radii vectores $r^0 = r(t^0), r = r(t)$ and the angle $\phi - \phi^0$ between them, if the energy h is fixed. In other words, $t - t^0$ is, for fixed h , a locally single-valued function of $r^0 = OP^0, r = OP$ and of the length $\rho = P^0P$ of the chord of the arc (P^0, P) . This holds, of course, not only in the Newtonian case $U = r^{-1}$.

But it will be shown that in the Newtonian case the time $t - t^0$ depends, for fixed h , on the sides r^0, r, ρ of the triangle PP^0O in such a way as to contain r^0 and r only in the combination $r^0 + r$; so that

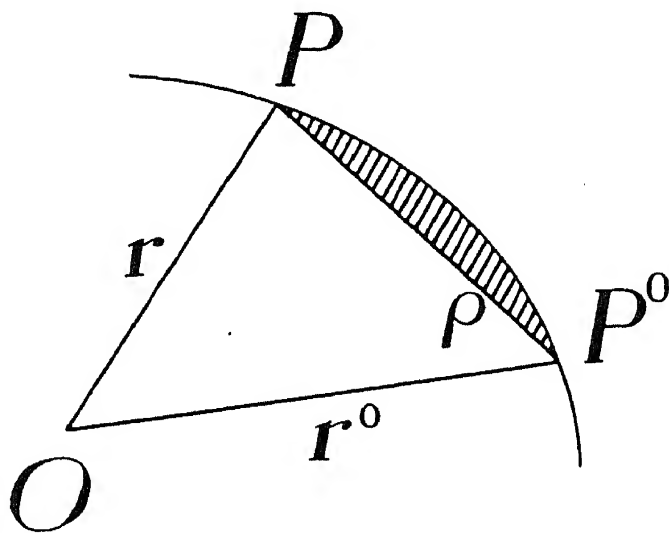


FIG. 4

$t - t^0$ is, for fixed h , a locally single valued function of the perimeter $r^0 + r + \rho$ and of the chord ρ . This theorem of Lambert, which is fundamental in the practice of determination of orbits, is by no means evident, since it does not hold for arbitrary laws $U(r) = r^{-\lambda}$ ($\lambda = \text{const.}$), for instance.

A proof of Lambert's theorem can be obtained by an application of the theorem of Gauss-Bonnet on the surface of revolution S_h of §244. However, the proof is shorter if use is made of the "Beltrami-Hilbert

integral" or the "isoenergetic action W " not via S_h but in a more direct manner, as follows:

§248. Since P^0 is fixed, one can consider the radius vector r and the chord ρ as bipolar coordinates of P , with O and P^0 as poles. Then (35), §56 (footnote) shows that the Lagrangian function (1), §241 takes, in terms of the coordinates $q_1 = \frac{1}{2}(r - \rho)$, $q_2 = \frac{1}{2}(r + \rho)$, the form

$$L = \frac{1}{2} \sum_{i=1}^2 \frac{(-1)^i (q_2 - q_1)}{q_i^2 - (\frac{1}{2}r^0)^2} q_i'^2 + \frac{1}{q_1 + q_2};$$

$$\text{hence, } H = \frac{1}{2} \sum_{i=1}^2 \frac{q_i^2 - (\frac{1}{2}r^0)^2}{(-1)^i (q_2^2 - q_1^2)} p_i^2 - \frac{1}{q_1 + q_2}$$

is the corresponding Hamiltonian function $H = H(p_1, p_2, q_1, q_2)$. Consequently, if $G(W_x, \chi)$ is, for fixed h , an abbreviation for

$$(11) \quad G(W_x, \chi) = -2\{\chi + h\chi^2\} + \{\chi^2 - (\frac{1}{2}r^0)^2\} W_x^2,$$

the partial differential equation (15), §114 is

$$(12) \quad H(W_{q_1}, W_{q_2}, q_1, q_2) = h, \quad \text{i.e., } G(W_{q_1}, q_1) = G(W_{q_2}, q_2).$$

Since (11) remains unchanged upon writing $-W_x$ for W_x , a solution $W = W(q_1, q_2)$ of the separated equation (12) may be obtained by integrating between $\chi = q_1$ and $\chi = q_2$ a certain function $f = f(\chi)$ of the single variable χ , this f being determined by the explicit form of G ; in fact, it is clear from (11) that $f(\chi)$ is the square root of $2\{(\chi + \frac{1}{2}r^0)^{-1} + h\}$. Thus, on introducing instead of χ the integration variable $\bar{r} = \chi + \frac{1}{2}r^0$, where $r^0 = \text{const.}$, one sees that the function

$$(13) \quad W = 2^{\frac{1}{2}} \int_{q_1 + \frac{1}{2}r^0}^{q_2 + \frac{1}{2}r^0} (\bar{r}^{-1} + h)^{\frac{1}{2}} d\bar{r};$$

$$(2q_1 = r - \rho, \quad 2q_2 = r + \rho; \quad ()^{\frac{1}{2}} \geq 0),$$

of the coordinates q_1, q_2 and of the integration constants h, r^0 satisfies (12). Hence, the last of the rules (22) of §116 bis implies* that

* Actually, §116 bis assumes that condition (18) of §116 is satisfied, which means, in the present case, that $W_{q_1 h} W_{q_2 r^0} - W_{q_1 r^0} W_{q_2 h} \neq 0$. But it is easily verified from the representation (13) of W that this condition is satisfied,

the partial derivative of (13) with respect to h is $t + \text{const.}$ But (13) shows that $W_h = 0$ if and only if $q_1 = q_2$, which means, in view of Fig. 4, that $r = r^0$, where r, r^0 belong to t, t^0 , respectively. Hence from (13),

$$(14) \quad t - t^0 = W_h \equiv 2^{-\frac{1}{2}} \int_{\frac{1}{2}(r^0+r-\rho)}^{\frac{1}{2}(r^0+r+\rho)} (\bar{r}^{-1} + h)^{-\frac{1}{2}} d\bar{r}; \quad ()^{-\frac{1}{2}} \geq 0.$$

Since the integral on the right of (14) depends, for a fixed energy h , only on the perimeter $r^0 + r + \rho$ and on the chord ρ , the proof of Lambert's theorem (§247) is complete.

§249. Clearly, the quadrature (14) leads to elementary functions of the integration limits. However, caution is necessary, since these elementary functions are not single-valued, (14) being an algebraic function with real branch points even in the simplest case, $h = 0$. Furthermore, (14) was derived on the assumption that t is sufficiently close to t^0 (cf. §246, as well as the fact that the rule $t - t^0 = W_h$ of §116 bis was proved by using the local existence theorem of differential equations). However, it is clear for reasons of analyticity that (14) becomes valid for arbitrary $t - t^0$ if, for every pair t, t^0 , one chooses a suitable branch of the elementary multivalued function defined by (14).

Excluding the rectilinear case ($c = 0, h \leq 0$), which can afterwards be included by an obvious limit process, one finds after a straightforward discussion that the correct choice of the respective branches leads to the following evaluation of (14):

In the parabolic case $h = 0$, the value of (14) is

$$(15_1) \quad t - t^0 = \frac{1}{6} \{ (r^0 + r + \rho)^{\frac{3}{2}} \mp (r^0 + r - \rho)^{\frac{3}{2}} \}, \quad (h = 0),$$

where the lower or the upper sign (i.e., $+$ or $-$) is valid according as the segment shaded in Fig. 4 does or does not contain the focus O .

In the hyperbolic case $h > 0$, define a unique pair u^0, u of real numbers by the conditions

$$(15_2') \quad \begin{aligned} u^0 &= 2 \operatorname{arc} \sinh \left\{ \frac{1}{2}(r^0 + r - \rho)h \right\}^{\frac{1}{2}}, \\ u &= 2 \operatorname{arc} \sinh \left\{ \frac{1}{2}(r^0 + r + \rho)h \right\}^{\frac{1}{2}}; \quad (0 < u^0 < u). \end{aligned}$$

Then the value of (14) is

unless the solution path is either rectilinear or circular. In these trivial cases, the validity of (14) easily follows either by a direct verification or by an obvious limit process.

$$(15_2) \quad t - t^0 = (2h)^{-\frac{3}{2}} \{ (\sinh u - u) \mp (\sinh u^0 - u^0) \}, \quad (h > 0),$$

where the lower or the upper sign is valid according as the segment shaded in Fig. 4 does or does not contain the focus O .

In the elliptic case $h < 0$, define a unique pair u^0, u of real numbers by the conditions

$$(15_3') \quad \begin{aligned} u^0 &= 2 \arcsin \left\{ -\frac{1}{2}(r^0 + r - \rho)h \right\}^{\frac{1}{2}}, \\ u &= 2 \arcsin \left\{ -\frac{1}{2}(r^0 + r + \rho)h \right\}^{\frac{1}{2}}; \quad (0 < u^0 < u < \pi), \end{aligned}$$

and suppose first that the segment shaded in Fig. 4 does not contain the empty focus F . Then the value of (14) is

$$(15_3) \quad t - t^0 = (-2h)^{-\frac{3}{2}} \{ (u - \sin u) \mp (u^0 - \sin u^0) \}, \quad (h < 0),$$

where the lower or the upper sign is valid according as the shaded segment does or does not contain the focus O . If, on the other hand, the shaded segment does contain F , then, without changing the definition of u^0 and u in (15₃') and the determination of the sign in (15₃), one has to

$$(15_3^*) \quad \text{replace } u \text{ in } (15_3) \text{ by } u^* = 2\pi - u.$$

It is understood that, in (15₁)–(15₃*), the root $A^{\frac{1}{2}}$ (or $A^{\pm\frac{1}{2}}$) of $A > 0$ is meant to be positive, and that t^0, t are defined in the way described in §246. Since the ambiguity mentioned in §246 arises, if $c \neq 0$, only in the elliptic case, it is quite natural that the rule (15₃')–(15₃*) belonging to $h < 0$ turns out to be more complicated than either of the rules (15₁) and (15₂)–(15₂').

From the formulae of §260 below, it is possible to check all the rules (15₁)–(15₃*).

§250. For a fixed h , let Σ_h denote the family of those solution paths in the (x, y) -plane which have the energy h . If $h < 0$, then (4) shows that Σ_h consists of those ellipses in the (x, y) -plane which have the origin O as a focus and $2a = -h^{-1}$ as common length of their major axes, while the eccentricity ($0 \leq e \leq 1$) and the direction of the major axis in the (x, y) -plane are arbitrary; so that every ellipse contained in Σ_h occurs in Σ_h in all possible positions about O . It is understood that the circle ($e = 0$) of diameter $-h^{-1}$ occurs only once, and that the radii of the circle of radius $-h^{-1}$ about O are considered as ellipses with eccentricity $e = 1$ ($e = 0$; cf. §243). A similar description of Σ_h is implied by (4) also when $h = 0$ or $h > 0$.

In what follows, it will be assumed that $h < 0$. Then the above

description of Σ_h is to the effect that Σ_h consists of all ellipses (incl. segments) which have the circle of radius $-h^{-1}$ about the common focus O as directrix. This directrix circle is, according to §243, the curve of zero velocity belonging to h , and will be denoted by D_h ; so that

$$(16) \quad D_h: \quad x^2 + y^2 = 4a^2, \quad \text{where} \quad a = (-2h)^{-1} > 0.$$

§251. Choose in the interior, but not at the centre O , of D_h a point P_0 , denote by $\Gamma_h(P_0)$ the subset of Σ_h consisting of those solution paths of energy h which go through P_0 , and let $E_h(P_0)$ be the ellipse which touches the circle D_h and has O and P_0 as foci. If $AB = BA$ denotes the distance between two points A, B , it is clear from (16) that the major axis of $E_h(P_0)$ is of length

$$(17) \quad (2a - OP_0) + OP_0 + (2a - OP_0) = 4a - OP_0 > 2a, \\ \text{since} \quad 0 < OP_0 < 2a.$$

Consequently, D_h is not a directrix of $E_h(P_0)$, and so $E_h(P_0)$ is not a solution path of energy h (the same remark holds, of course, in the excluded case $P_0 = O$ also, since $E_h(O) = D_h$). Actually, $E_h(P_0)$ is the envelope† of the solution paths which constitute the subset $\Gamma_h(P_0)$ of Σ_h .

In what follows, it will be necessary to consider on any ellipse C contained in $\Gamma_h(P_0)$, the points $P_0^* = P_0^*(C)$ and $P_0^{**} = P_0^{**}(C)$ in-

† In fact, the equations characterizing the points P of an ellipse C contained in $\Gamma_h(P_0)$ and the points Q of $E_h(P_0)$ are

$$(I) \quad C: \quad OP + PF = 2a; \quad (II) \quad E_h(P_0): \quad OQ + QP_0 = 4a - OP_0,$$

$2a$ and (17) being the lengths of the major axes, while O, F and O, P_0 are the foci, of C and $E_h(P_0)$, respectively. Since P_0 and P_0^* are on C ,

$$(III) \quad OP_0 + P_0F = 2a; \quad (IV) \quad OP_0^* + P_0^*F = 2a; \quad (V) \quad P_0F + P_0^*F = P_0P_0^*,$$

(V) being clear from Fig. 5. If P is any point of the ellipse C such that $P \neq P_0^*$, then Fig. 5 shows that either the three points P_0, P, F are not collinear or $P = P_0$; so that $P_0P < P_0F + PF$ in both cases. This inequality, when combined with (I) and (III), can be written as $OP + PP_0 < 4a - OP_0$; so that P is, in view of (II), in the interior of the ellipse $E_h(P_0)$. If, on the other hand, $P = P_0^*$, then (III), (IV), (V) show that (II) is satisfied by $Q = P_0^*$.

Accordingly, a point P of C is within or on $E_h(P_0)$ according as $P \neq P_0^*$ or $P = P_0^*$. But there is, for a fixed P_0 , only one P_0^* (cf. Fig. 5). Hence, the ellipses C and $E_h(P_0)$ touch each other at their common point P_0^* . Since this holds for any ellipse C contained in $\Gamma_h(P_0)$, it follows that $E_h(P_0)$ is the envelope of $\Gamma_h(P_0)$.

licated in Fig. 5, where $F = F(C)$ denotes the empty focus of the path C (in the sense of §246). It is understood that $P_0^* \neq P_0 \neq P_0^{**}$ also when P_0 is collinear with the foci O, F of C , in which case $P_0^* = P_0^{**}$.

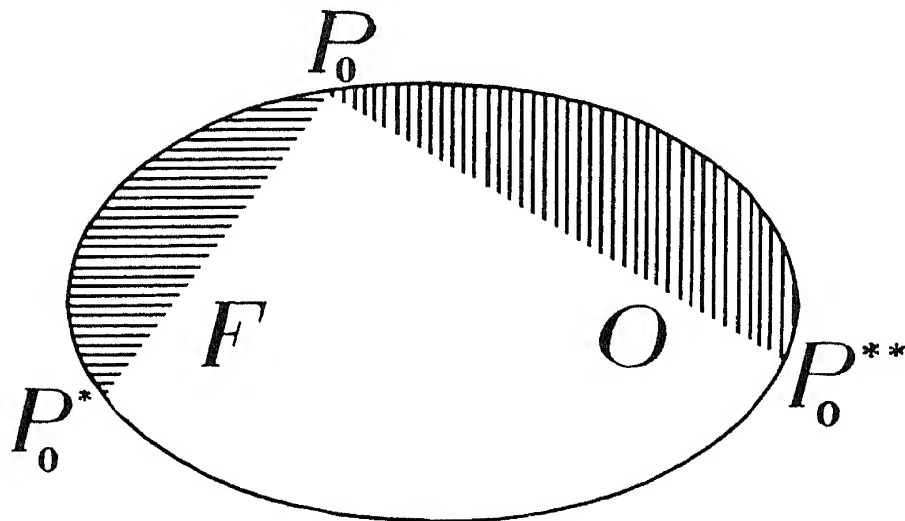


FIG. 5

§252. The situation becomes more intuitive if one takes a slightly different start, as follows:

Let there be given, besides $P_0 (\neq O)$, a point $P (\neq O)$ within the circle D_h . If R is either P_0 or P , let B_R denote the circle which touches D_h and has the centre R ; so that the radius of B_R is $2a - RO$, by (16). Since the solution paths which constitute Σ_h have, by §250, the common focus O and the common directrix D_h , it is clear that a given solution path of energy $h (< 0)$ goes through both points P_0, P if and only if its empty focus (§246) is a common point of the two circles B_{P_0}, B_P . It is also clear that B_{P_0} and B_P intersect at two distinct points, touch each other or do not meet according as P lies within, on or without the ellipse which touches D_h and has P_0 as a focus. Since this ellipse is the ellipse $E_h(P_0)$ defined at the beginning of §251, it follows that the number of solution paths of energy h which go through both points P_0, P is 2, 1 or 0 according as P is within, on or without $E_h(P_0)$; cf. Fig. 6. (This clearly implies that $E_h(P_0)$ is the envelope of $\Gamma_h(P_0)$; cf. §251.)

If P_0 is fixed and P is chosen within or on $E_h(P_0)$, let $C = C_P(P_0)$ and $C' = C_{P'}(P_0)$ be the two solution paths of energy h which go through both points P, P_0 ; so that $C_P \neq C_{P'}$ or $C_P = C_{P'}$ according as P is within or on $E_h(P_0)$. In either case, let $F = F_P(P_0)$ and $F' = F_{P'}(P_0)$ denote the empty foci (§246) of C and C' , respectively, while O is a focus of both C and C' . Finally, let $I = I_P = I_{P'}(P_0)$

denote the common chord $[P_0, P]$ of C and C' ; so that I is the major axis in the limiting case $C = C'$. It is easy to see that, if P is within $E_h(P_0)$, i.e., if $C \neq C'$, then one of the two ellipses C, C' , say C , has both of its foci, O and F , on the same side of the chord I of C ; while the foci, O and F' , of C' are separated from each other by the chord I of C' .

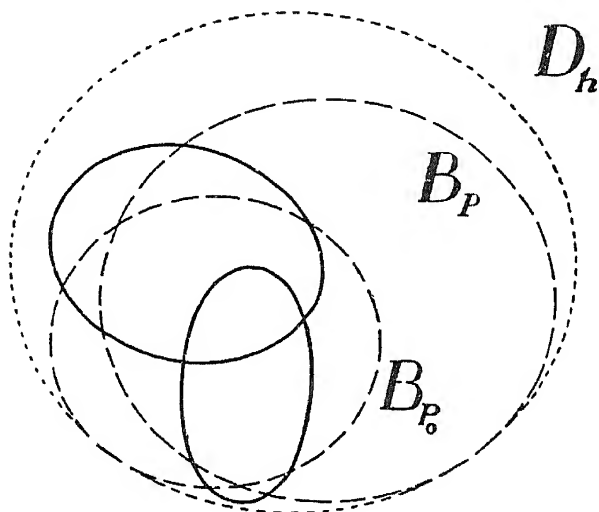


FIG. 6

§253. The preceding elementary considerations enable one to discuss the problem of minima with respect to the homogeneous calculus of variations problem $\bar{\delta}W = 0$, where

$$(18) \quad W = \int_{P_0}^P \{2(U + h)(x'^2 + y'^2)\}^{\frac{1}{2}} d\bar{l};$$

$$U = r^{-1}, \quad r = (x^2 + y^2)^{\frac{1}{2}},$$

h has a preassigned value which is, for the present, negative, and the dash of $\bar{\delta}$ means that the boundary points P_0, P are not varied when the δ -process is applied to W (cf. §95 and §172).

First, the integrand of (18) is the function (11), §179 which belongs to the present problem, (1)–(2₁), §241. Hence, the end of §172 states that the set Σ_h of the solution arcs of energy h is identical with the set of the regular (i.e., unbroken) extremal arcs of $\bar{\delta}W = 0$. Finally, §177 shows that the question, whether a given solution arc P_0P of energy h does or does not yield a minimum of W , is identical with the question concerning the location of conjugate points. Now, §250–§252 supply the answer to this question in the elliptic case $h < 0$. In fact, on comparing the end of §252 with the fact that $E_h(P_0)$ is, by §251–§252, the envelope of those solution paths of en-

ergy h which go through P_0 , one clearly† arrives at the following result:

In order that the elliptic extremal arc which joins P with P_0 yield a proper strong minimum of (18), one has to choose this arc on the ellipse C , and not on the ellipse C' , if the notations are the same as at the end of §252. Furthermore, if P_0 is fixed and P varies on this C , then that arc (P_0, P) of C which does not contain P_0^{**} will represent the strong minimum as long as P lies between P_0 and P_0^* when the positive orientation of C is that leading from P_0 via P_0^* to P_0^{**} ; cf. Fig. 5. When P passes from the left to the right of the conjugate point P_0^* of P_0 in Fig. 5, the positively oriented arc (P_0, P) of C ceases to represent a minimum (even a weak minimum). Finally, the limit, $P = P_0^*$, between proper strong minimum and no minimum at all requires a direct discussion, and corresponds to the coalescence of the two ellipses C, C' , i.e., to the case where P is on $E_h(P_0)$; cf. §252.

Barring this limiting case and uniting the consideration of the two extremal ellipses C, C' which have an arc (P_0, P) , one sees that there are altogether four cases possible, according as the elliptic segment which is determined by the oriented arc (P_0, P) of the extremal contains neither, both, one or the other of the two foci. These four cases are identical with the four cases which one obtains by combining the alternative (\mp) of (15_3) with the permutation (15_3^*) .

§254. It is assumed in §253 that P_0 and P can be joined by at least one (and, of course, at most two) solution arcs of energy h . According to §252, this is the case if and only if P lies within or on $E_h(P_0)$. Hence, the problem $\bar{\delta}W = 0$ does not possess any regular extremal if P is chosen in the exterior of $E_h(P_0)$. In this case, a proper strong minimum of (18) is furnished by a broken extremal P_0Q_0QP which, in Fig. 7, is represented by the portions $[P_0, Q_0]$, $[Q, P]$ of the radii $[O, Q_0]$, $[O, Q]$ together with the arc (Q_0, Q) of

† In virtue of the envelope construction of conjugate points in calculus of variations.

It should be mentioned that the particular problem $\bar{\delta}W = 0$ defined by (18), where $h < 0$, was the first example discussed by Jacobi when he introduced the theory of conjugate points. The name "conjugate point" in calculus of variations originated precisely from this example, since in this example these points are the points which are conjugate points in the sense of the theory of conics; cf. Fig. 5.

Similarly, the broken extremal of Fig. 7, pointed out by Todhunter, is perhaps the earliest instance of a discontinuous solution of a regular problem in calculus of variations.

the directrix (16), i.e., the curve of zero velocity. This is shown by verifying that the standard sufficient conditions for an extremal

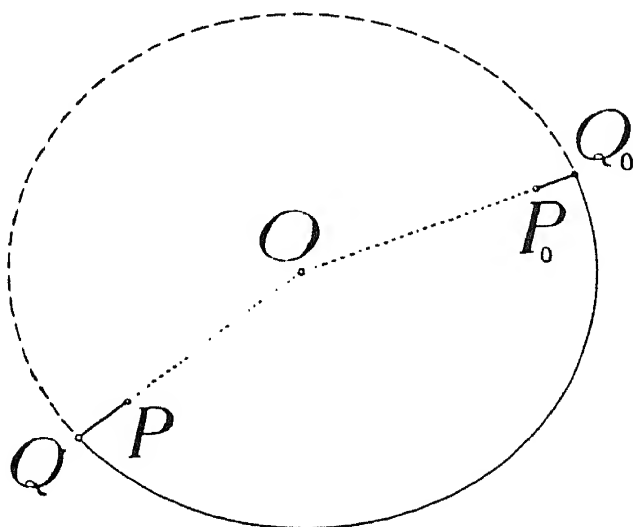


FIG. 7

which yields a proper strong minimum are satisfied, at least if O is not collinear with the points P_0, P (which are, of course, to be chosen within D_h). This broken extremal exists also when P is within or on $E_h(P_0)$, i.e., also when the regular extremals of §253 exist.

Notice that the portions $[P_0, Q_0]$, $[Q, P]$ of radii of D_h are, by §243, regular extremals of (18); that, at the points Q_0, Q of the broken extremal of Fig. 7, the well-known corner condition of transversality is satisfied; finally, that the circular arc (Q_0, Q) does not contribute to (18), since, by (2₂), the factor $(U + h) \equiv (r^{-1} + h)$ of $(x'^2 + y'^2)$ in (18) vanishes along this arc.

§255. Only the elliptic case $h < 0$ has been considered thus far. If $h \geq 0$, one might expect that, in view of the last remark of §244, the solution path of energy h yields a proper strong minimum of (18) for arbitrarily distant P_0, P on this path. This is, however, not true, since it will turn out that, just as in the rule of §253, one has to choose between two conics even when $h \geq 0$. It will be seen that the actual simplification arising for $h \geq 0$ is to the effect that the case of §254 in which one cannot join P_0, P by at least one solution path of energy h is possible only when $h < 0$. Correspondingly, there does not exist a curve of zero velocity for $h \geq 0$ (cf. §243 and §254).

§256. Consider first the case $h > 0$. According to §242, the family, Σ_h , of all solution paths of energy h consists of those hyperbolas which have the origin, O , of the (x, y) -plane as a focus and possess a transverse axis of length $-2a = h^{-1} > 0$; while the direction of this

axis and the eccentricity are arbitrary. It is understood that by an hyperbola of focus O is meant that branch of the hyperbola which turns its concavity towards O , and that the half-lines issuing from O must be included as hyperbolas of minimum eccentricity, $e = 1$. The considerations of §252 can be adapted to the present case $h > 0$, as follows:

Let there be given two distinct points P_0, P in the (x, y) -plane such that $P_0 \neq O \neq P$. If R is either P_0 or P , let B_R denote the circle which has the point R as centre and $-2a + OR$ as radius. Since $-2a = h^{-1} > 0$, the sum of the two radii vectores $-2a + OP_0$, $-2a + OP$ exceeds the distance P_0P ; so that the circles B_{P_0}, B_P always intersect at two distinct points, say at F and F' . It follows, therefore, from the definition of an hyperbola and from the above description of the family Σ_h , that there exist, for every pair P_0, P , exactly two solution paths of energy h , say C and C' , which join P_0 with P ; the empty foci of C and C' being F and F' , respectively. It is also seen that if I denotes the common chord, $[P_0, P]$, of C and C' , and if one excludes the limiting case where O, P_0, P are collinear, then I intersects the transverse axis of one of the hyperbolas, say that of C , between the foci, O and F , of C ; while the intersection of I with the transverse axis of C' takes place beyond O , i.e., in such a way that the foci, O and F' , of C' are situated on the same side of I .

Hence, on adapting from §252–§253 the construction of the solution paths of energy h which go through P_0 , one sees that P_0 has no conjugate point in the interior of the arc (P_0, P) of C , while either an interior point or the end point P of the arc (P_0, P) of C' is a conjugate point of P_0 on C' ; the conjugate point being situated in the interior of the arc (P_0, P) or at P according as the common chord, $I = [P_0, P]$, of C and C' does not or does go through the common focus, O , of C and C' .

Accordingly, the two points P_0, P of the (x, y) -plane can always be joined by an arc, (P_0, P) , of a solution path, C , of given energy $h > 0$ in such a way as to yield for the integral (18) a proper strong minimum; while the arc (P_0, P) of C' represents not even a weak minimum (at least if the chord $I = [P_0, P]$ does not go through O ; this limiting case corresponds to that case in §252–§253 in which P lies on $E_h(P_0)$ and requires a direct discussion).

Since I intersects the transverse axes of C and of C' on the side of the vertex and beyond O , respectively, the two solution arcs (P_0, P) , (P_0, P) of the extremal problem correspond to the alternative sign in (15₂).

§257. The remaining case, $h = 0$, may be thought of as a limiting case either of the complicated elliptic case (§253) or of the hyperbolic case (§256). It is, however, preferable to proceed in a direct manner, as follows:

According to §242, the family, Σ_0 , of all solution paths of energy $h = 0$ consists of those parabolas which have the origin, O , of the (x, y) -plane as focus and possess as directrix a line which has an arbitrary distance, p , from O , and an arbitrary direction. It is understood that the half-lines issuing from O must be included as parabolas of parameter $p = 0$.

For two given points P_0, P of the (x, y) -plane which are distinct from O , let B_{P_0}, B_P denote the circles which go through O and respectively have P_0, P as centre. Let T and T' denote the two common tangents of B_{P_0} and B_P , and let N and N' denote the lines through O which are perpendicular to T and T' , respectively. It is clear from the definition of a parabola and from the above description of Σ_0 , that N and N' and only these lines are axes, and T and T' directrices, of parabolas which are solution paths through both points P, P_0 . If C and C' denote these two parabolas, then $C = C'$ if and only if $T = T'$, which means that O, P_0, P are collinear. Excluding this limiting case, one sees that the common chord, $I = [P_0, P]$, of C and C' cuts the axis of one of the two parabolas, say the axis N of C , on the same side of the focus O as the directrix T of C ; while the axis N' of C' is cut by I beyond the common focus O . The balance of the necessary considerations, as well as the final result, is the same as in the hyperbolic case of §256.

Clearly, the alternative sign in (15₁) corresponds to the two possibilities which are represented by the arcs $(P_0, P), (P, P_0)$ of the respective parabolas C, C' .

The Anomalies

§258. According to §241, the integration of $[L]_x = 0, [L]_y = 0$, i.e., of

$$(1_1) \quad x'' + xr^{-3} = 0, \quad y'' + yr^{-3} = 0;$$

$$(1_2) \quad L = \frac{1}{2}(x'^2 + y'^2) - r^{-1}; \quad (r^2 = x^2 + y^2),$$

when considered as a problem concerning *orbits* (loci) in the (x, y) -plane which are *not* referred to a time parameter, can be obtained from the integrals

$$(2_1) \quad \frac{1}{2}(x'^2 + y'^2) - r^{-1} = h; \quad (2_2) \quad xy' - yx' = c$$

without any real quadrature; cf. §218 bis. This fact, fundamental in the practise of determination of orbits, does not hold in case the Newtonian law $U = r^{-1}$ is replaced by an arbitrary law, say of the form $U = r^{-\lambda}$. If an orbit is known in terms of its geometrical integration constants (4), §241, the time elapsed between two given positions on the orbit is supplied by the formulae of §249.

All of this dodges, however, the question of the general solution of (1₁). In fact, obtaining the coordinates x, y as functions of the time for a given set of integration constants requires awkward elimination processes between the formulae of §241 and §249. For this reason, x, y will now be treated directly as functions of a time variable and of the integration constants.

§259. To this end, one can apply the transformation of §230, by choosing $z = z(\zeta)$ to be $z = \zeta^2$; so that $|z_\zeta|^2 = 4(\xi^2 + \eta^2)$,

$$(3_1) \quad x = \xi^2 - \eta^2, \quad y = 2\xi\eta; \quad (3_2) \quad \dot{t} = 4(\xi^2 + \eta^2) = 4(x^2 + y^2)^{\frac{1}{2}} = 4r,$$

where the dot denotes differentiation with respect to the new time variable, \bar{t} . Since $|z_\zeta|^2/r = 4$, comparison of (12₁)–(12₂), §230, with (1₂), §258, gives $\bar{U} = 4 + 4(\xi^2 + \eta^2)h$ and $\bar{\omega} \equiv 0$; so that, from (13₁)–(13₂), §230,

$$(4_1) \quad \xi = 8h\dot{\xi}, \quad \eta = 8h\dot{\eta}; \quad (4_2) \quad -(\dot{\xi}^2 + \dot{\eta}^2) + 8(\xi^2 + \eta^2)h = -8.$$

The coordinates ξ, η are the parabolic coordinates of §54, with $r = 0$ as origin. On placing

$$(5_1) \quad \gamma = (\mp 32h)^{\frac{1}{2}} \text{ if } h \leq 0 \text{ and } \gamma = 4 \text{ if } h = 0;$$

$$(5_2) \quad u = \gamma\bar{t}; \quad (h \leq 0, \gamma > 0),$$

and denoting by α, β integration constants, one sees that (4₁) is satisfied* by

$$(6) \quad \begin{aligned} \xi &= \alpha \cos \tfrac{1}{2}u, & \eta &= \beta \sin \tfrac{1}{2}u; \\ \xi &= \alpha \cosh \tfrac{1}{2}u, & \eta &= \beta \sinh \tfrac{1}{2}u; \\ \xi &= \tfrac{1}{2}\alpha, & \eta &= \tfrac{1}{2}\beta u, \end{aligned}$$

where $h < 0, h > 0, h = 0$, respectively. However, (6) must satisfy

* No generality is lost by replacing the four integration constants of (4₁) by the two integration constants α, β chosen in (6). This will become clear, in §261, by comparison of the results with §242, and is explained by the fact that one can choose arbitrarily both the direction of the x -axis in the (x, y) -plane and the origin of the t -axis.

the invariant relation (4₂). Since $\cos^2 u = 1 - \sin^2 u$, $\cosh^2 u = 1 + \sinh^2 u$, one sees from (5₁)–(5₂) that (4₂) subjects the constants α, β of (6) to the condition $\alpha^2 \pm \beta^2 = -h^{-1}$ if $h \leq 0$, and to $\beta^2 = 2$ if $h = 0$. This means that if $h \leq 0$, and if one puts $a = (-2h)^{-1} \geq 0$, then there exists a unique $e \geq 0$ such that $\alpha^2 = a(1 - e)$, $\beta^2 = \pm a(1 + e)$; while if $h = 0$, then $\alpha^2 = 2p$, $\beta^2 = 2$ for a unique $p \geq 0$.

§260. Substituting (6) into (3₁) and using the representation of α, β just obtained, one easily finds that

$$\begin{aligned} x &= a(\cos u - e), \quad y = a\sqrt{(1 - e^2)} \sin u, \quad \text{if } h < 0, \\ &\hspace{25em} (a > 0, 0 \leq e \leq 1); \\ (7) \quad x &= a(\cosh u - e), \quad y = a\sqrt{(e^2 - 1)} \sinh u, \quad \text{if } h > 0, \\ &\hspace{25em} (a < 0, e \geq 1); \\ x &= \frac{1}{2}(p - u^2), \quad y = (\sqrt{p})u, \quad \text{if } h = 0, \hspace{10em} (p \geq 0). \end{aligned}$$

Since $r^2 = x^2 + y^2$, it follows that, according as $h < 0$, $h > 0$ or $h = 0$,

$$(8) \quad r = a(1 - e \cos u), \quad -a(e \cosh u - 1), \quad \frac{1}{2}(p + u^2).$$

This implies that if t_0 denotes an integration constant,

$$(9) \quad t - t_0 = (\sqrt{a^3})(u - e \sin u), \quad (\sqrt{-a^3})(e \sinh u - u), \quad (\sqrt{\frac{1}{4}})(pu + \frac{1}{3}u^3)$$

in the three respective cases. In fact, it is seen from (3₂) and (5₂), where $\dot{t} = dt/d\bar{t}$ and $\gamma = \text{const.}$, that $t = 4\gamma^{-1} \int r du$. Hence, (9) follows from (8), (5₁) and the definition, $a = (-2h)^{-1}$, of $a \geq 0$ for $h \leq 0$.

§261. Let, for a moment,

$$(10) \quad \bar{x} = x + ae \text{ if } h \leq 0, \text{ and } \bar{x} = x - \frac{1}{2}p \text{ if } h = 0; \text{ while } \bar{y} = y \text{ for } h \geq 0,$$

and put

$$(11) \quad \begin{aligned} b^2 &= \pm a^2(1 - e^2) \text{ if } h \leq 0; \\ \text{so that } b^2 &> 0 \text{ if } e \neq 1, \text{ and } b^2 = 0 \text{ if } e = 1, \end{aligned}$$

by (7). Clearly, (7) can be written, for $h < 0$; $h > 0$; $h = 0$, as

$$(12) \quad \begin{aligned} \bar{x} &= a \cos u, & \bar{y} &= b \sin u; \\ \bar{x} &= a \cosh u, & \bar{y} &= b \sinh u; \\ \bar{x} &= -\frac{1}{2}u^2, & \bar{y} &= (\sqrt{p})u. \end{aligned}$$

But (12) represents an ellipse, an hyperbola and a parabola in the respective cases; the centre and the length of the axes being $(\bar{x}, \bar{y}) = (0, 0)$ and $\pm 2a > 0$, $2|b| \geq 0$ in the first two cases, while $(\bar{x}, \bar{y}) = (0, 0)$ is the vertex and $p \geq 0$ the parameter in the third case. It follows, therefore, from (10) that $(x, y) = (0, 0)$ is a focus in all three cases; while (11) shows that e is the eccentricity in the first two cases. Hence, the constants $a, e; p$ occurring in (7) are identical with the constants $a, e; p$ of §242; so that, by (4), §241,

$$(13) \quad \begin{aligned} e &= (1 + 2hc^2)^{\frac{1}{2}} \geq 0, \quad a = (-2h)^{-1} \text{ if } h \neq 0; \text{ and} \\ p &= c^2 \geq 0 \text{ if } h = 0. \end{aligned}$$

The geometrical meaning of the auxiliary time variable u , which is connected with t by (9), is clear from (10)–(12).

The auxiliary time variable u is called the eccentric anomaly.

§262. Notice that the geometrical meaning of u is lost if $c = 0$, where $h \leq 0$, i.e., if either $e = 1$, where $h \leq 0$, or $p = 0$, where $h = 0$. If $c \neq 0$, the motion is direct or retrograde about $(x, y) = (0, 0)$ according as $c > 0$ or $c < 0$; cf. §214. This ambiguity is represented by the square roots occurring in (7), (9), and, implicitly, in (11)–(12). It is easy to verify that these square roots are to be chosen so as to have the same sign as c (in particular, the minor axis, $2b$, turns out to be negative for $h < 0$, if the motion is retrograde). In view of §214, one can assume without loss of generality that the motion is direct ($c > 0$, $\sqrt{} > 0$; $h \leq 0$), provided that it is defined what is direct, i.e., provided that the motion is not rectilinear ($c = 0$, $\sqrt{} = 0$; $h \leq 0$).

§263. Suppose that the motion is not rectilinear, i.e., that $c \neq 0$. Then $0 \leq e < 1$, $e > 1$ or $p > 0$ according as $h < 0$, $h > 0$ or $h = 0$. Hence, if r, w denote polar coordinates about the focus $(x, y) = (0, 0)$, and $w = 0$ belongs to the periastron, i.e., to the minimum of the radius vector, the equation of the conic is

$$(14) \quad \begin{aligned} r &= \frac{a(1 - e^2)}{1 + e \cos w} \text{ if } (-2a)^{-1} = h \leq 0, \text{ and} \\ r &= \frac{p}{1 + \cos w} \text{ if } h = 0. \end{aligned}$$

According to (8), the minimum of r is attained at $u = 0$ in all three cases $h \leq 0$; and $u = 0$ implies, by (9), that $t = t_0$. Hence, if ϕ is

the polar angle with reference to the positive half of the x -axis, and if it is not assumed that the x -axis is an axis of the conic, then

$$(15_1) \ x = r \cos \phi, \ y = r \sin \phi; \quad (15_2) \ w = \phi - \omega; \quad (15_3) \ \omega = (\phi)_{t=t_0},$$

where the integration constant (15_3) is undetermined mod 2π and can, except in the circular case ($e = 0$), be characterized also by the property that $\min r(\phi) = r(\omega)$.

The variable (15_2) which occurs in (14) is called the true anomaly.

§264. If $c > 0$, the true anomaly w and the eccentric anomaly u are steadily increasing functions of each other. Furthermore, with the positive determination of the square roots,

$$(16) \quad \begin{aligned} \tan \frac{1}{2}w &= \sqrt{\{(1+e):(1-e)\}} \tan \frac{1}{2}u, \\ \sqrt{\{(e+1):(e-1)\}} \tanh \frac{1}{2}u, \quad u/\sqrt{p} \end{aligned}$$

in the respective cases $h < 0$, $h > 0$, $h = 0$. In fact, substitution of (15_1) – (15_3) into (2_1) gives $r^2 w' = c$; so that $w' > 0$, and so $w = w(t)$ is a steadily increasing function of t . The same holds, in view of (3_2) and (5_2) , for $u = u(t)$. Furthermore, $u = 0$ belongs, by (8) and §263, to $w = 0$. Hence, it is easily verified from the first of the two identities

$$(17) \quad \begin{aligned} \cos \alpha &= (1 - \tan^2 \frac{1}{2}\alpha)/(1 + \tan^2 \frac{1}{2}\alpha), \\ \sin \alpha &= (2 \tan \frac{1}{2}\alpha)/(1 + \tan^2 \frac{1}{2}\alpha), \end{aligned}$$

that (8) and (14) imply (16), with $\sqrt{\ } > 0$.

§265. If $c \neq 0$, it is convenient to introduce still another time variable, ζ , by placing

$$(18) \quad \zeta = n(t - t_0), \quad \text{where} \quad n^2 = a^{-3}, \quad -a^{-3}, \quad 4p^{-3}$$

in the respective cases $h < 0$, $h > 0$, $h = 0$, and where the integration constant n is chosen to be positive or negative according as $c > 0$ or $c < 0$. The integration constant t_0 occurring in (18) is meant to be the same as the one which is implicitly defined by (14) – (15_3) .

For reasons which in §276 will become obvious for $h < 0$, one calls the linear function (18) of t the mean anomaly.

§265 bis. It is clear from (9) and (18) that the eccentric anomaly $u = u(t)$ coincides with the mean anomaly $\zeta = \zeta(t)$ only in the circular case $e = 0$, in which case $u = u(t)$ is, in view of (16), identical with the true anomaly $w = w(t)$ also.

Generally, the function

$$(19) \quad \varepsilon = w - \zeta$$

of t is called the equation of the centre.†

§266. The relation (9) between the time and the eccentric anomaly is what is called (at least if $h < 0$ and $e < 1$) the equation of Kepler. The importance of this equation is clear from the fact that, in order to obtain $x = x(t)$, $y = y(t)$ from (7), one has to know u as a function of t , i.e., one has to solve Kepler's equation (9) with respect to u . If $c \neq 0$, one can write (9) with the help of (18) in the form

$$(20) \quad \zeta = u - e \sin u, \quad e \sinh u - u, \quad (u + \frac{1}{3}u^3/p)/\sqrt{p}.$$

§267. According to §214, one can obtain $r = r(t)$, for every fixed value of the integration constant $xy' - yx' = c$, from a problem $[L^*]_r = 0$ with a single degree of freedom, to which §185–§190 are applicable.

Let, in particular, $h < 0$, and exclude the limiting case $e = 1$, where $c = 0$, and also the case $e = 0$, where $r(t) = \text{const.}$ Then §188 is applicable to the periodic solutions $r = r(t)$ of $[L^*]_r = 0$. Actually, the uniformizing time variable, \bar{t} , which belongs to $[L^*]_r = 0$ in virtue of §188, is precisely the eccentric anomaly, u , of the elliptic motion. In other words, (7), §188 holds for $q = r$, $\bar{t} = u$, if β, α denote the maximum and the minimum of r . This is readily verified from §188 and (8)–(9), §260, the extrema of r being $\alpha = a(1 - e)$ and $\beta = a(1 + e)$. Correspondingly, (9₁), §188 is represented by (9), §260, if one chooses $t_0 = 0$; so that $v_0 = \sqrt{a^3}$, $\lambda_1 = -e\sqrt{a^3}$, while $\lambda_n = 0$ for $n > 1$. Finally, the Fourier coefficients (10₂), §188, where $q(\bar{t}) \equiv r(u)$, lead to Bessel functions; cf. §278 below.

The correspondence between r and t needs a uniformization not only in the elliptic case just described but also in the hyperbolic and parabolic cases ($h \geq 0$, $c \neq 0$). In fact, in these cases $r(t)$ attains every value greater than $\min r(t) = r_0 > 0$ exactly twice, when t runs from $-\infty$ to $+\infty$. According to (8) and (9), the eccentric anomaly is a uniformizing variable in these cases also.

† The origin of this name is explained by the remark that centuries ago the astronomers used the words "equation" and "inequality" for what to-day one would call "correction" and "deviation," respectively. Thus, "equation of the centre" means something like "correction for the deviation from circular motion." Now, (19) is a measure for this deviation, $w(t) = \zeta(t)$ being true only in the circular case $e = 0$.

Actually, the eccentric anomaly u is, according to (7) and (9), a uniformizing variable not only for r and t but also for (x, y) and t , no matter what are the values of the integration constants h, c . This is the analytical significance of the eccentric anomaly.

§268. In §263–§267, but not in §260, it is assumed that $c \neq 0$. Now let $c = 0$; so that the motion is, by §242, rectilinear, and can, therefore, be assumed to take place along the x -axis. Then $y(t) \equiv 0$; so that $r \equiv (x^2 + y^2)^{\frac{1}{2}} = |x|$, and so (1₁), (2₁) reduce to

$$(21_1) \quad x'' + x|x|^{-3} = 0; \quad (21_2) \quad \frac{1}{2}x'^2 = |x|^{-1} + h.$$

Since the mass which rests at the origin attracts the moving particle with a force which increases when the distance $|x|$ decreases, it is readily shown without an explicit integration of (21₁), that every solution $x = x(t)$ of the problem (21₁) with a single degree of freedom must tend to zero when t tends to a suitable finite t_0 ; so that no motion is possible without a collision of the two particles. Furthermore, the algebraic differential equation (21₁) has a singularity at $x = 0$; and, what is more, every solution $x = x(t)$ of (21₁) becomes singular at $t = t_0$, if $x(t) \rightarrow 0$ as $t \rightarrow t_0$. In fact, (21₂) shows that $|x'| \rightarrow \infty$ as $|x| \rightarrow 0$.

It will be shown that the eccentric anomaly is a local regularizing variable of this singularity of the analytic function $x(t)$ of t .

First, $c = 0$ is, according to (13), characterized by $e = 1$ in the elliptic and hyperbolic cases $h \leq 0$, and by $p = 0$ in the parabolic case $h = 0$. Hence, (7) and (9) reduce to

$$(22_1) \quad x = a(\cos u - 1); \quad a(\cosh u - 1); \quad -\frac{1}{2}u^2;$$

$$(22_2) \quad t - t_0 = (\sqrt{a^3})(u - \sin u); \quad (\sqrt{-a^3})(\sinh u - u); \quad u^3/\sqrt{36}.$$

Accordingly, there is a collision for $u = 0, 2\pi, \dots$ or only for $u = 0$ according as $h < 0$ or $h \geq 0$. For reasons of periodicity, it is sufficient to consider $u = 0$ alone also when $h < 0$.

Choosing the origin of the t -axis so that $t = 0$ corresponds to $u = 0$, i.e., that $t_0 = 0$, one can write (22₁)–(22₂) in all three cases as $x = u^2 P_1(u)$, $t = u^3 P_2(u)$, where $P_j(z)$ is, for $j = 1, 2$, a power series which converges for all z , has real coefficients, and a non-vanishing constant term $P_j(0)$. It follows, therefore, by local elimination of u at $u = 0$ ($t = 0$) that, for sufficiently small $|t|$,

$$(23) \quad x(t) = (\sqrt[3]{t})^2 \sum_{n=0}^{\infty} c_n (\sqrt[3]{t})^n, \text{ where } c_0 \neq 0 \text{ and } c_n \geq 0.$$

This implies that $x(t)$ is of the same sign for small positive as for small negative t , i.e., that the particle which moves on the x -axis is reflected through the collision by the particle which rests at $x = 0$. In other words, the situation is the same as in §170, only that the path is now rejected at a state at which the velocity $x'(t)$ is infinite, instead of being zero; in fact, $|x'(t)|$ is, by (23), of the order $|t|^{-\frac{1}{2}}$ at $t = 0$.

§269. The precise description of this situation is, however, as follows:

Consider u and t as complex variables (which are eventually restricted to be real). Then $x = x(t)$ is an analytic function of t , since $x(t)$ is obtained by elimination of u between the entire functions (22_1) – (22_2) of u . According to (23), the analytic function $x(t)$ has at $t = 0$ an algebraic branch point at which three sheets of the Riemann surface unite. It also follows from (23) that, if $t \neq 0$ is real and small, $x(t)$ is real on exactly one of the three sheets, whether $t \rightarrow -0$ or $t \rightarrow +0$, i.e., whether the state is before or after the collision. Accordingly, if $0 \neq t \rightarrow \pm 0$, then $x(t)$ acquires a singularity through which exactly one *real* analytic continuation is possible.

This unique real branch of the analytic continuation can be considered as defining a dynamical continuation of the problem. It is clear from the rather non-analytic implications of the last remark of §268 that this continuation, as well as the result of §268 concerning the rejection of the moving particle by the collision, is such that an observatory situated on either of the particles will hardly be in a position to issue a bulletin on observations made during the collision, or, what is the same thing, rejection. On the other hand, these considerations have a clear analytical meaning, and describe the real singularities of the problem, i.e., those singularities of the analytic function $x(t)$ of the complex variable t which lie at real t when only real-valued branches of $x(t)$ are considered.

§270. Since (22_1) – (22_2) is a parametrization of (23), the eccentric anomaly u not only uniformizes the multivalued relationship between (x, y) and t or r and t (§267), but it also uniformizes, in all three cases $h \gtrless 0$ of a rectilinear motion ($c = 0$), the local singularities of the real analytic function $x(t)$ of the real variable t (§269).

The second, but not the first, of these uniformizations will turn out to be possible in case of more than two bodies also, provided that only two of the bodies collide. While no explicit formulae (7)–(9)

will then be available, a local uniformizing variable, u , will be, as it is in (5₁)–(5₂) by (3₂), such that $t = t(u)$ becomes a linear function of the integral of the reciprocal value of the vanishing distance $r = r(u)$; cf. §414, §448, §498.

§271. The results of §268–§269 are by no means evident, and are indeed wrong if one replaces the Newtonian by an arbitrary law of attraction. In fact, suppose that the attraction is inversely proportional to the third, instead of the second, power of the distance. Then $|x|^{-3}$ in (21₂) must be replaced by x^{-4} ; so that $\frac{1}{2}(dx/dt)^2 = x^{-2} + h$. Hence, $t = t(x)$ follows by an elementary quadrature, which, when inverted, shows that $x = x(t)$ has at the moment, say $t = 0$, of collision ($x = 0$) a logarithmical singularity if $h \neq 0$, while $x = \sqrt{(8t)}$ if $h = 0$. Now, in the first case no analytic, in the second case no *real* analytic, continuation of $x = x(t)$ is possible through $t = 0$; so that, for two different reasons, the results of §268–§269 do not hold in either case.

There is a further difference between the Newtonian case $U(r) = r^{-1}$ and the present case $U(r) = r^{-2}$. For if $U = r^{-1}$, then $c = 0$ is not only sufficient but, in view of §242, also necessary for a collision. On the other hand, it is easily verified from (16₂)–(16₃), §214 that if $U = r^{-2}$, there can be a collision also when $c \neq 0$. Cf. also §162, §374 bis.

§272. In the parabolic case $h = 0$, one has from (17), (16), (9)

$$(24_1) \quad \cos w = \frac{p - u^2}{p + u^2}, \quad \sin w = \frac{2(\sqrt{p})u}{p + u^2};$$

$$(24_2) \quad \tan \frac{1}{2}w + \frac{1}{3} \tan^3 \frac{1}{2}w = \frac{2(t - t_0)}{\sqrt{p^3}},$$

if $c \neq 0$. The cubic equation (24₂) is equivalent to the case $h = 0$ of (20) and is, since Halley's work (1705) on his comet, fundamental in the practice of determination of orbits.

According to (18), one can write (24₂) as $z + \frac{1}{3}z^3 = \zeta$, where $z = \tan \frac{1}{2}w$. Hence, if t is considered as a complex variable, $z = u/\sqrt{p}$ is a three-valued algebraic function of $\zeta = n(t - t_0)$. Since the zeros of the first derivative of $\zeta = \zeta(z) \equiv z + \frac{1}{3}z^3$ are at $z = \pm i$, points at which the second derivative does not vanish, only two of the three sheets of the Riemann surface of $z = z(\zeta)$ unite at either of the two finite branch points $\zeta = \pm i + \frac{1}{3}(\pm i)^3 = \pm \frac{2}{3}i$; while all three sheets unite at $\zeta = \infty$. Since $z = z(\zeta)$ has no further

finite singularities, it follows that if $\zeta_* = n(t_* - t_0)$ is any fixed *real* number, that branch of $z = u/\sqrt{p}$ which is real for real $\zeta = n(t - t_0)$ may be developed according to the powers of $\zeta - \zeta_*$ into a power series, with $|\pm \frac{2}{3}i - \zeta_*| = (\frac{4}{9} + \zeta_*^2)^{\frac{1}{2}}$ as radius of convergence. Thus, although there are no real singularities, the radius of convergence is finite for every ζ_* , and varies with ζ_* so as to attain at $\zeta_* = 0$ its minimum, $\frac{2}{3}$.

§273. If $h > 0$, then (7), (8), (9), (16) contain hyperbolic functions and are, therefore, inconvenient from the point of view of the computer. This technical inconvenience can, however, be avoided by rewriting the formulae belonging to $h > 0$ in such a way that their numerical treatment becomes possible by using *real* trigonometric and logarithmic tables only. To this end, one has merely to replace the eccentric anomaly, u , by another real time variable, $u = u(u)$, which is usually referred to as Lambert's angle and is defined by $\tan \frac{1}{2}u = \tanh \frac{1}{2}u$. In fact, this may be written, on the one hand, as $u = \log \tan \frac{1}{2}(u + \frac{1}{2}\pi)$ and implies, on the other hand, that $\cosh u = \sec u$, $\sinh u = \tan u$, by (17). Hence, the transition from u to $u = u(u)$ requires only trigonometric and logarithmic tables on the one hand, and it removes from (7), (8), (9), (16) the hyperbolic functions on the other hand.

Expansions of the Elliptic Motion Into Fourier Series

§274. In what follows, only the elliptic case $h < 0$ will be considered. It will sometimes be necessary to exclude the limiting case $e = 1$ of periodic collisions (§268–§269) and the trivial case $e = 0$ of circular motions. Assuming, without loss of generality (§242), that $c \geq 0$, and placing

$$(1_1) \quad f \equiv f(e) = \frac{e}{1 + (1 - e^2)^{\frac{1}{2}}} \equiv \frac{1 - (1 - e^2)^{\frac{1}{2}}}{e};$$

$$(1_2) \quad g \equiv g(e) = \frac{e \exp (1 - e^2)^{\frac{1}{2}}}{1 + (1 - e^2)^{\frac{1}{2}}},$$

where the square roots are positive, one has, if $0 \neq e \neq 1$,

$$(2) \quad 0 < f(e) < e < g(e) < 1.$$

The last inequality is easily verified by showing that the derivative of (1₂) with respect to e is positive for $0 < e < 1$; so that

$$(3) \quad 0 = g(0) < g(e_1) < g(e_2) < g(1) = 1, \quad \text{if} \quad 0 < e_1 < e_2 < 1.$$

§275. According to the formulae (18), (15₂)–(15₃), (9) of the preceding section, one has

$$(4_1) \quad n = a^{-\frac{3}{2}}; \quad (4_2) \quad (\zeta)_{t=t_0} = 0; \quad (4_3) \quad (w)_{t=t_0} = 0; \quad (4_4) \quad (u)_{t=t_0} = 0;$$

while (7)–(8) reduce to

$$(5_1) \quad x = a(\cos u - e), \quad y = a(1 - e^2)^{\frac{1}{2}} \sin u; \quad (5_2) \quad r = a(1 - e \cos u),$$

and (14)–(15₁) to

$$(6_1) \quad \begin{aligned} x &= r \cos (w + \omega) \\ y &= r \sin (w + \omega); \end{aligned} \quad (6_2) \quad r = \frac{a(1 - e^2)}{1 + e \cos w}; \quad (6_3) \quad \omega = \text{const.},$$

where, according to (16), (18), (20),

$$(7_1) \quad \tan \frac{1}{2}w = \left(\frac{1 + e}{1 - e} \right)^{\frac{1}{2}} \tan \frac{1}{2}u; \quad (7_2) \quad \zeta = n(t - t_0);$$

$$(7_3) \quad \zeta = u - e \sin u.$$

Application of (17), §264 either to $\alpha = w$ or to $\alpha = u$ gives

$$(8) \quad \cos w = \frac{-e + \cos u}{1 - e \cos u}, \quad \sin w = \frac{(1 - e^2)^{\frac{1}{2}} \sin u}{1 - e \cos u}$$

if use is made of (7₁). The inversion of (8) is

$$(9) \quad \cos u = \frac{e + \cos w}{1 + e \cos w}, \quad \sin u = \frac{(1 - e^2)^{\frac{1}{2}} \sin w}{1 + e \cos w},$$

since (7₁) remains unchanged if one replaces u by w and e by $-e$. It also follows from (7₁) that

$$(10) \quad \cos \frac{1}{2}w = \frac{(1 - e)^{\frac{1}{2}} \cos \frac{1}{2}u}{(1 - e \cos u)^{\frac{1}{2}}}, \quad \sin \frac{1}{2}w = \frac{(1 + e)^{\frac{1}{2}} \sin \frac{1}{2}u}{(1 - e \cos u)^{\frac{1}{2}}},$$

the square roots being again positive, by §264 ($c > 0$). Replacing u by w and e by $-e$, one sees that the inversion of (10) is

$$(11) \quad \cos \frac{1}{2}u = \frac{(1 + e)^{\frac{1}{2}} \cos \frac{1}{2}w}{(1 + e \cos w)^{\frac{1}{2}}}, \quad \sin \frac{1}{2}u = \frac{(1 - e)^{\frac{1}{2}} \sin \frac{1}{2}w}{(1 + e \cos w)^{\frac{1}{2}}}.$$

According to (5₂) or (6₂), one can write (10) or (11) as

$$(12) \quad r^{\frac{1}{2}} \cos \frac{1}{2}w = (1 - e)^{\frac{1}{2}} \cos \frac{1}{2}u, \quad r^{\frac{1}{2}} \sin \frac{1}{2}w = (1 + e)^{\frac{1}{2}} \sin \frac{1}{2}u.$$

Finally, it is easily verified from (5₂) that (8) may be written as

$$(12 \text{ bis}) \quad \begin{aligned} r \cos (w-u) &= a \left\{ 1 - e \cos u - \frac{1}{2} [1 + P(e^2)] e^2 \sin^2 u \right\}, \\ r \sin (w-u) &= a \left\{ 1 - \frac{1}{2} [1 + P(e^2)] e \cos u \right\} e \sin u, \end{aligned} \quad (P(0)=0),$$

if one puts $1 - (1 - e^2)^{\frac{1}{2}} = \frac{1}{2}e^2[1 + P(e^2)]$; so that $P(e^2) = \frac{1}{4}e^2 + \dots$ is an even power series which converges for $|e| < 1$ and vanishes at $e = 0$.

It should be mentioned that, according to (1₁),

$$(13_1) \quad (1 \pm e \cos w)f = \frac{1}{2}(1 \pm 2f \cos w + f^2)e; \quad (13_2) \quad e = 2f/(1 + f^2).$$

§276. The connection between the time t and the three anomalies ζ , u , w can be defined by the initial conditions (4₂)–(4₄) and the quadratures which are assigned by

$$(14_1) \quad \frac{dt}{d\zeta} = a^{\frac{3}{2}}; \quad (14_2) \quad \frac{d\zeta}{du} = \frac{r}{a}; \quad (14_3) \quad \frac{du}{dw} = \frac{r}{a(1 - e^2)^{\frac{1}{2}}}.$$

In fact, (14₁) is clear from (7₂), (4₁). Similarly, (14₂) follows from (7₃), (5₂). Finally, on differentiating the first of the relations (9) with respect to w and then using the second and (6₂), one obtains (14₃).

The name “mean anomaly” is derived from the fact that $\zeta = \zeta(t)$ would be the true anomaly $w = w(t)$ if the angular velocity $w' = w'(t)$

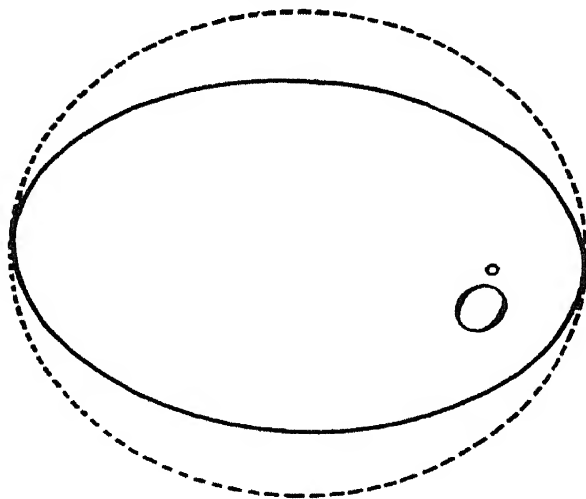


FIG. 8

about the origin of the plane of the Cartesian coordinates (6₁) were independent of t . Actually, on writing the integration constant (4₁) in the form

$$(15) \quad n = 2\pi:T, \quad \text{where} \quad T^2 = 4\pi^2 a^3, \quad (a = -\frac{1}{2}h^{-1}),$$

one sees from (7₂), (7₃) and (5₁) that T is the period of the elliptic motion $x = x(t)$, $y = y(t)$. But T is, in view of the relation (15) which expresses the third law of Kepler, independent of the eccentricity, i.e., the amount of time needed for a complete circuit about the focus is the same for $0 < e < 1$ (and even for $e = 1$) as for $e = 0$, if the length of the major axis is fixed. Finally, it is clear that in the circular case $e = 0$ the constant angular velocity $w'(t)$ becomes the constant n .

Since the three anomalies ζ , u , w are, in view of (14₁)–(14₃), strictly increasing functions of t or of one another, one can use any of them as time variable. The period with reference to t being T , it is seen from (5₁)–(7₃) and (15) that the period with reference to any of the three anomalies is 2π .

§277. In particular, any of the (analytic) functions $u - \zeta$, x , r , $\cos w$, \dots of t , when considered as a function $F = F(\zeta)$ of $\zeta = n(t - t_0)$, can be developed into a Fourier series

$$(16_1) \quad F(\zeta) = \sum_{k=-\infty}^{+\infty} A_k \exp(k\zeta i);$$

$$(16_2) \quad A_k = \frac{1}{2\pi} \int_0^{2\pi} F(\zeta) \exp(-k\zeta i) d\zeta.$$

These A_k lead to the transcendental entire functions

$$(17_1) \quad J_m(z) = \frac{1}{2\pi} \int_0^{2\pi} \cos(mu - z \sin u) du \equiv (-1)^m J_{-m}(z);$$

$$(17_2) \quad J_m(z) = \sum_{n=0}^{\infty} \frac{(-1)^n (\frac{1}{2}z)^{m+2n}}{n!(m+n)!},$$

($m = 0, 1, \dots$) which satisfy the recursion formulae

$$(18_1) \quad J_{k-1}(z) + J_{k+1}(z) = 2kJ_k(z)/z;$$

$$(18_2) \quad J_{k-1}(z) - J_{k+1}(z) = 2dJ_k(z)/dz,$$

and, though usually associated with the name of Bessel, have been used extensively, precisely in this connection (which is that of Bessel), and more than half a century prior to Bessel, by Lagrange and others.*

* In addition, early investigations on boundary value problems (D. Bernoulli, Euler; Fourier, Poisson) had also led to these functions.

§278. First, from (5₂), (7₃) and (16₂),

$$(19) \quad A_k = \frac{1}{2\pi} \int_0^{2\pi} (1 - e \cos u) F(u - e \sin u) \exp(-kiu + kei \sin u) du.$$

Choose, in particular, $F(\zeta) = \exp l u i$, where l is a fixed positive integer and $u = u(\zeta)$. Then a partial integration of (19) shows that, in view of the definition (17₁), one has $A_k = J_{k-l}(ke)l/k$ if $k \neq 0$, while $A_0 = -\frac{1}{2}e$ or $A_0 = 0$ according as $l = 1$ or $l > 1$. In other words, (16₁) reduces for $F(\zeta) = \exp l u i$ to*

$$(20) \quad \frac{\cos lu}{l} = \sum_{k=-\infty}^{+\infty} \frac{J_{k-l}(ke)}{k} \cos k\zeta, \quad \frac{\sin lu}{l} = \sum_{k=-\infty}^{+\infty} \frac{J_{k-l}(ke)}{k} \sin k\zeta$$

if $l = 2, 3, \dots$;

while if $l = 1$, then

$$(21) \quad \cos u = -\frac{1}{2}e + \sum_{k=-\infty}^{+\infty} \frac{J_{k-1}(ke)}{k} \cos k\zeta, \quad \sin u = \sum_{k=-\infty}^{+\infty} \frac{J_{k-1}(ke)}{k} \sin k\zeta.$$

Substitution of (21) into (7₃) and (5₂) gives

$$(22_1) \quad u = \zeta + e \sum_{k=-\infty}^{+\infty} \frac{J_{k-1}(ke)}{k} \sin k\zeta;$$

$$(22_2) \quad \frac{r}{a} = 1 + \frac{1}{2}e^2 - e \sum_{k=-\infty}^{+\infty} \frac{J_{k-1}(ke)}{k} \cos k\zeta.$$

On differentiating (22₁) with respect to ζ , one sees from (14₂) that

$$(23) \quad \frac{a}{r} = 1 + e \sum_{k=-\infty}^{+\infty} J_{k-1}(ke) \cos k\zeta \equiv 1 + 2 \sum_{k=1}^{\infty} J_k(ke) \cos k\zeta,$$

by (18₁). Similarly, on expressing $\cos u$, $\cos 2u$ from (21), (20), and noting that (5₂) gives $r^2 = a^2(1 + \frac{1}{2}e^2 - 2e \cos u + \frac{1}{2}e^2 \cos 2u)$, one obtains

$$(24) \quad \frac{r^2}{a^2} = 1 + \frac{3}{2}e^2 - 4 \sum_{k=1}^{\infty} \frac{J_k(ke)}{k^2} \cos k\zeta.$$

Furthermore, from (23) and (6₂),

* The prime in \sum' means that $k \neq 0$.

$$\begin{aligned}
 (25) \quad \cos w &= -e + (1 - e^2) \sum_{k=-\infty}^{+\infty} J_{k-1}(ke) \cos k\zeta; \quad \text{and} \\
 \sin w &= (1 - e^2)^{\frac{1}{2}} \sum_{k=-\infty}^{+\infty} J_{k-1}(ke) \sin k\zeta
 \end{aligned}$$

follows by differentiating (22₂) with respect to ζ , since

$$\sin w = \frac{(1 - e^2)^{\frac{1}{2}}}{ae} \frac{dr}{d\zeta} \quad (\text{in fact, } \sin u = \frac{r \sin w}{a(1 - e^2)^{\frac{1}{2}}} \text{ and } \frac{dr}{d\zeta} = ae \sin u \cdot \frac{a}{r},$$

by (9), (6₂) and (5₂), (14₂), respectively). Also, from (5₁) and (21),

$$\begin{aligned}
 (26) \quad x &= -\frac{3}{2}ae + a \sum_{k=-\infty}^{+\infty} \frac{J_{k-1}(ke)}{k} \cos k\zeta, \\
 y &= a(1 - e^2)^{\frac{1}{2}} \sum_{k=-\infty}^{+\infty} \frac{J_{k-1}(ke)}{k} \sin k\zeta.
 \end{aligned}$$

Differentiating (26) twice with respect to ζ , then using (14₁), (6₁)–(6₂), and comparing the result with (1₁), §258, one obtains

$$\begin{aligned}
 (27) \quad \frac{a^2 \cos(w + \omega)}{r^2} &= \sum_{k=-\infty}^{+\infty} k J_{k-1}(ke) \cos k\zeta, \\
 \frac{a^2 \sin(w + \omega)}{r^2} &= \sum_{k=-\infty}^{+\infty} k J_{k-1}(ke) \sin k\zeta.
 \end{aligned}$$

This procedure can be continued indefinitely. The above expansions are those occurring most often in the applications of the theory of perturbations to the solar system.

§279. According to (7₂), the formulae (26) represent the Fourier expansions of the Cartesian coordinates $x = x(t)$, $y = y(t)$. The corresponding results (25), (22₂) for the polar coordinates $r = r(t)$, $w = w(t)$ [cf. (6₁)–(6₃)] are less complete, since (25) corresponds only to (21), while the analogue to (22₁) is missing. There exists a Fourier series

$$(28) \quad w = \zeta + \sum_{k=1}^{\infty} C_k(e) \sin k\zeta$$

which corresponds to (22₁), but the integral which defines the Fourier constant $C_k(e)$ will turn out to be a new transcendent,

$$(29) \quad C_m(z) = \frac{\sqrt{(1-z^2)}}{\pi m} \int_0^{2\pi} \frac{\cos(mu - zm \sin u)}{1 - z \cos u} du, \\ (|z| < 1; m = 1, 2, \dots),$$

where the square root is $+1$ at $z = 0$. The (even) function (29) of z is regular analytic in the circle $|z| < 1$ but not at $z = 1$, while (17₁)–(17₂) is a transcendental entire function; hence, (28) is of a more advanced type than any of the Fourier series of §278. However, $C_k(z)$ can be expressed as an infinite series of Bessel functions (17₁)–(17₂), as follows:

$$(30) \quad C_k(z) = \frac{2}{k} \sum_{n=-\infty}^{+\infty} \frac{z^{|n|} J_{k+n}(kz)}{[1 + \sqrt{(1-z^2)}]^{|n|}}, \\ (|z| < 1; k = 1, 2, \dots).$$

In fact, if $|z| < 1$ and

$$(31) \quad f = \frac{z}{1 + \sqrt{(1-z^2)}}, \text{ then} \\ \frac{\sqrt{(1-z^2)}}{1 - z \cos u} \equiv \frac{1 - f^2}{1 - 2f \cos u + f^2} = \sum_{n=-\infty}^{+\infty} f^{|n|} \cos nu,$$

the last expansion, where $|f| < 1$, being standard*; while $|z| < 1$ readily implies that $|f| < 1$. Thus, on inserting (31) into (29) and then using (17₁), one obtains (30).

In order to prove (28)–(29), notice first that the difference $w - \zeta$ is, in view of the formulae of §275, an odd function of ζ , and has the period 2π . Thus, (28) is the Fourier series of $w - \zeta$, the Fourier constants being

$$C_k(e) = \frac{1}{\pi} \int_0^{2\pi} (w - \zeta) \sin k\zeta d\zeta;$$

so that

$$\pi k C_k(e) = \int_0^{2\pi} \cos k\zeta d(w - \zeta),$$

by partial integration. Since $\int_0^{2\pi} \cos k\zeta d\zeta = 0$, it follows that

* A proof follows, for instance, by differentiating with respect to ψ the identity (37₂) below.

$$\begin{aligned}\pi k C_k(e) &= \int_0^{2\pi} \cos k\zeta dw = \int_0^{2\pi} \cos k\zeta \frac{dw}{du} du \\ &= \int_0^{2\pi} \cos k\zeta \frac{a(1-e^2)^{\frac{1}{2}}}{a(1-e \cos u)} du,\end{aligned}$$

by (14₃) and (5₂). Hence, (29) follows from (7₃).

§280. Using the notation (19), §265 bis for the equation of centre, one can write (28), (30) also as

$$(32_1) \quad \varepsilon = w - \zeta; \quad (32_2) \quad \varepsilon = \sum_{k=1}^{\infty} C_k(e) \sin k\zeta;$$

$$(32_3) \quad C_k(e) = 2k^{-1} \sum_{n=-\infty}^{+\infty} f^{|n|} J_{k+n}(ke),$$

$f = f(e)$ being the function (1₁). From (32₂)-(32₃),

$$(33) \quad \varepsilon = 2 \sum_{k=1}^{\infty} \left(\sum_{j=-\infty}^{+\infty} J_j(ke) f^{|k-j|} / k \right) \sin k\zeta, \text{ where } j \neq 0; f = f(e),$$

[cf. (1₁)].

Differentiating (33) and (32₁) with respect to ζ , one obtains

$$(34) \quad (1-e^2)^{\frac{1}{2}} a^2/r^2 = 1 + 2 \sum_{k=1}^{\infty} \left(\sum_{j=-\infty}^{+\infty} J_j(ke) f^{|k-j|} \right) \cos k\zeta,$$

by (14₂)-(14₃).

§281. The explicit expansions of the three anomalies u , w , $\zeta = n(t - t_0)$ in terms of one another are as follows: Corresponding to the elementary inversion (7₂) of (22₁), the inversion of (28) is the elementary expansion

$$(35) \quad \zeta = w + 2 \sum_{k=1}^{\infty} \frac{1 + k(1-e^2)^{\frac{1}{2}}}{(-1)^k k} f^k \sin kw; f = f(e), \quad [\text{cf. (1}_1)],$$

and also the remaining pair $w = w(u)$, $u = u(w)$ is elementary,

$$(36_1) \quad w = u + 2 \sum_{k=1}^{\infty} \frac{f^k}{k} \sin ku; \quad (36_2) \quad u = w + 2 \sum_{k=1}^{\infty} \frac{(-f)^k}{k} \sin kw.$$

The series (35)-(36₂) for $\zeta - w$, $w - u$ belong to the oldest instances of Fourier series (Clairaut, d'Alembert, Euler), and can be verified as follows:

Separating in $-\log(1 - z) = \sum_{k=1}^{\infty} z^k/k$ reals and imaginaries, one obtains

$$(37_1) \quad -\frac{1}{2} \log(1 - 2\rho \cos \psi + \rho^2) = \sum_{k=1}^{\infty} \frac{\rho^k}{k} \cos k\psi;$$

$$(37_2) \quad \arctan \frac{\rho \sin \psi}{1 - \rho \cos \psi} = \sum_{k=1}^{\infty} \frac{\rho^k}{k} \sin k\psi,$$

where $z = \rho \exp i\psi$, $\rho = |z| < 1$. Thus, logarithmic differentiation of (13₁) gives

$$(38) \quad \frac{e \sin w}{1 - e \cos w} = \frac{d}{dw} \log(1 - 2f \cos w + f^2) = 2 \sum_{k=1}^{\infty} f^k \sin kw,$$

where (37₁) has been applied to $\psi = w$, $\rho = f < 1$; cf. (2). On replacing w in (38) by $w + \pi$ and then applying the second of the relations (9), one obtains

$$(39) \quad \begin{aligned} -e(1 - e^2)^{-\frac{1}{2}} \sin u &= 2 \sum_{k=1}^{\infty} (-f)^k \sin kw; \text{ hence,} \\ \zeta &= u + 2(1 - e^2)^{\frac{1}{2}} \sum_{k=1}^{\infty} (-f)^k \sin kw, \end{aligned}$$

by (7₃), and so (35) is equivalent to (36₂).

On the other hand, (36₂) is equivalent to (36₁). In fact, (36₁) and (36₂) go over into each other if one replaces u by w and f by $-f$; while (2) shows that f goes over into $-f$ if e is replaced by $-e$. But (7₁) remains unchanged if one replaces u by w and e by $-e$.

Accordingly, it is sufficient to prove (36₁). Now, from (7₁),

$$(40) \quad \tan \frac{w - u}{2} = \frac{f \sin u}{1 - f \cos u}, \text{ since } f = \frac{(1 + e)^{\frac{1}{2}} - (1 - e)^{\frac{1}{2}}}{(1 + e)^{\frac{1}{2}} + (1 - e)^{\frac{1}{2}}},$$

by (1₁). Finally, comparison of (40) with (37₂) proves (36₁), the connection being $\rho = f$, $\psi = u$.

§282. As a consequence, (25) and (21) have the elementary analogues

$$(41) \quad \begin{aligned} \cos w &= -f + (1 - f^2) \sum_{k=1}^{\infty} f^{k-1} \cos ku, \\ \sin w &= (1 - f^2) \sum_{k=1}^{\infty} f^{k-1} \sin ku, \end{aligned}$$

or, conversely,

$$(42) \quad \cos u = f + (1 - f^2) \sum_{k=1}^{\infty} (-f)^{k-1} \cos kw,$$

$$\sin u = (1 - f^2) \sum_{k=1}^{\infty} (-f)^{k-1} \sin kw.$$

In fact, it is seen from (13₂) that the first of the relations (39) is identical with the second of the relations (42), and that the first of the relations (41) is, in view of (6₂), equivalent to

$$(43) \quad \frac{a}{r} = \frac{1+f^2}{1-f^2} \left(1 + 2 \sum_{k=1}^{\infty} f^k \cos ku \right), \quad \frac{1+f^2}{1-f^2} = \frac{1}{(1-e^2)^{\frac{1}{2}}}, \quad \text{by (13}_2\text{)}.$$

Since (41) and (42) go over into each other if one interchanges w , f and u , $-f$, it follows that it is sufficient to verify (43). But (43) is clear from (14₃) by differentiation of (36₁).

§283. If $c_k = c_k(e)$, where $k = 0, \pm 1, \pm 2, \dots$, denotes the k -th Fourier coefficient in any of the Fourier series of §279–§282, then, since the periodic functions developed are regular analytic functions of the respective real variables $\zeta = n(t - t_0)$, u , w , the convergence of the Fourier series is so strong that $|c_k| < \vartheta^{|k|}$ for a suitable $\vartheta = \vartheta(e)$ which is less than 1. Excluding the circular case $e = 0$ (in which case $c_k = 0$ for all sufficiently large $|k|$), one can even obtain for the Fourier constants $c_k = c_k(e)$ an explicit asymptotic formula in terms of the k -th powers either of $f = f(e)$ or of $g = g(e)$, numbers which satisfy the inequalities (2). This asymptotic formula is clear from (35)–(36₂), (41)–(43) in the case of $f = f(e)$; while (20)–(27) and (28) belong to $g = g(e)$, since, if e is fixed ($0 < e < 1$) and $m \rightarrow +\infty$, then the functions (17₁) and (29) satisfy the relations

$$(44_1) \quad J_m(me) \sim \frac{1}{(1-e^2)^{\frac{1}{2}}} \frac{(g(e))^m}{(2\pi m)^{\frac{1}{2}}}; \quad (44_2) \quad C_m(e) \sim \frac{(g(e))^m}{m}.$$

In fact, by an asymptotic formula which was first established (Carlini, Jacobi; Cauchy) precisely in this connection and is, to-day, standard,

$$(44_1 \text{ bis}) \quad J_m(m \operatorname{sech} \alpha) \sim (2\pi m \tanh \alpha)^{-\frac{1}{2}} \exp \{ (\tanh \alpha - \alpha)m \} \\ \text{as } m \rightarrow +\infty,$$

where $\alpha > 0$ is arbitrarily fixed. Since there exists for every positive $e < 1$ exactly one positive $\alpha = \alpha(e)$ satisfying $1/e = \cosh \alpha \equiv 1/\operatorname{sech} \alpha$, one sees from (1₂) that (44₁ bis) may be written in the form (41₁). On the other hand, the formula (41₂), which is not implied by (41₁) and (30), follows from (29) by the same method as (44₁ bis) or (44₁) does from (17₁), namely, by Cauchy's method of "steepest descent," as rediscovered by Riemann. This method shows also that, in the excluded case $e = 1$ of periodic collisions, one has to replace (44₁), (44₂) by

$$(45_1) \quad J_m(m) \sim \frac{6^{\frac{1}{2}} \Gamma(\frac{1}{3})}{3^{\frac{1}{2}} m^{\frac{1}{2}} \pi}; \quad (45_2) \quad \tilde{C}_m(1) \sim -\frac{6^{\frac{1}{2}} \Gamma(\frac{2}{3})}{3^{\frac{1}{2}} m^{\frac{1}{2}} \pi},$$

where $\tilde{C}_m(1)$ denotes the limit† of $C_m(e)/(1 - e^2)^{\frac{1}{2}}$ as $e \rightarrow 1 - 0$.

§284. Writing z for e , one sees from (44₁) that $|J_m(mz)|^{1/m}$ has, as $m \rightarrow +\infty$, the limit $|g(z)|$. It is known from the theory of Bessel functions that this limit relation holds not only for $0 < e = z < 1$ but also for all imaginary z , i.e., for $z = i|z|$; so that

$$(46_1) \quad \lim_{m \rightarrow +\infty} |J_m(im|z|)|^{1/m} = |g(i|z|)|;$$

$$(46_2) \quad |g(i|z|)| = \frac{|z| \exp(1 + |z|^2)^{\frac{1}{2}}}{1 + (1 + |z|^2)^{\frac{1}{2}}},$$

(46₂) being implied by the definition (1₂) of g . Also

$$(47_1) \quad |g(i|z_1|)| < |g(i|z_2|)| \quad \text{if} \quad |z_1| < |z_2|;$$

$$(47_2) \quad (-i)^m J_m(i|z|) = \sum_{n=0}^{\infty} \frac{|\frac{1}{2}z|^{m+2n}}{n!(m+n)!}.$$

In fact, logarithmical differentiation of (46₂) shows that the derivative of $|g(i|z|)|$ with respect to $|z|$ is everywhere positive. This implies (47₁); while (47₂) is clear from (17₂).

According to (47₁), the function (46₂) is steadily increasing with $|z|$ from $|g(0)| = 0$ to $|g(+\infty i)| = +\infty$. This implies that the transcendental equation $|g(i\rho^*)| = 1$ has exactly one positive root ρ^* , and that for this unique ρ^* and for every $|z|$ one has

$$(48) \quad |g(i|z|)| \leq 1 \quad \text{according as} \quad |z| \leq \rho^*.$$

† It is understood that the integral (29), which is divergent at $z = 1$, can be defined at $z = 1$ either as a principal value or as a complex integral in which the integration path is deformed so as to avoid the poles.

Substitution of $|z| = \frac{2}{3}$ into (46₂) shows that the number $|g(\frac{2}{3}i)|$ exceeds 1 by a very small amount. This means, by (48), that ρ^* is somewhat less than $0.666 \dots$. Actually, ρ^* is somewhat greater than 0.66, since (46₂) shows that 1 exceeds $|g(0.66i)|$. The first decimals of ρ^* are found to be

$$(49) \quad \rho^* = 0.6627434 \dots$$

Expansions According to Powers of the Eccentricity

§285. According to §266, Kepler's problem requires, in the elliptic case $0 < e < 1$ under consideration, the determination of the solution $u = u(e; \zeta)$ of the transcendental equation (7₃). In order to obtain an expansion of the function $u = u(e; \zeta)$ which is implicitly defined by Kepler's equation (7₃), one can choose between two reasonable possibilities:

(i) In view of §278, one can develop, for every fixed value of the positive eccentricity $e (< 1)$, the deviation of the eccentric anomaly $u = u(e; \zeta)$ from the mean anomaly ζ into a Fourier series which proceeds according to trigonometric functions of the multiples of the variable ζ , and has coefficients which depend on the fixed value of the eccentricity.

(ii) On the other hand, one can also attempt to develop, for every fixed value of the mean anomaly ζ , the solution $u = u(e; \zeta)$ of Kepler's equation (7₃) into a Taylor series which proceeds according to powers of the variable eccentricity e , and has coefficients which depend on the fixed value of ζ ; so that

$$(50) \quad u = \sum_{j=0}^{\infty} c_j(\zeta) \frac{e^j}{j!}, \quad \text{where} \quad c_j = c_j(\zeta) = \left(\frac{\partial^j u(e; \zeta)}{\partial e^j} \right)_{e=0}.$$

§286. The expansion mentioned under (i) is given by (22₁), and can be written, in view of (18₁), as

$$(51) \quad u = \zeta + 2 \sum_{m=1}^{\infty} \frac{J_m(me)}{m} \sin m\zeta,$$

since $(-1)^m J_{-m}(z) = J_m(z) = (-1)^m J_m(-z)$, by (17₁), (17₂). It is clear from (44₁) and (3), or, more directly, from the elementary theory of Fourier series, that, no matter how close is the fixed value of the eccentricity $e (< 1)$ to 1, the series (51) is uniformly convergent for $-\infty < \zeta < +\infty$.

The problem concerning the expansion mentioned under (ii) is

much more involved. In fact, this expansion, namely (50), is a power series in e , and so the question of its convergence, in contrast with the convergence of (51), depends on an investigation of the singularities which the analytic function $u = u(e; \zeta)$ of e may exhibit for *complex* values of e , when ζ has an arbitrarily fixed real value (this is the reason that the relations of §284 will be needed for complex values of $z = e$ also). Furthermore, the coefficients of the power series (50) in e depend on ζ . Thus, if ρ denotes the radius of convergence of (50), then ρ is a function $\rho(\zeta)$ of the real angular variable ζ ; and it turns out that $\rho = \rho(\zeta)$ is not independent of ζ . Incidentally, it is sufficient to study the function $\rho(\zeta)$ for $0 \leq \zeta \leq \frac{1}{2}\pi$, since, the motion being symmetric with respect to both Cartesian coordinate axes, one clearly has

$$(52) \quad \rho(\zeta) = \rho(\zeta + \pi) = \rho(-\zeta); \quad (-\infty < \zeta < +\infty).$$

It will be shown in §287–§288 that

$$(53_1) \quad \rho(\zeta) \geq \rho^* \text{ for } -\infty < \zeta < +\infty; \quad (53_2) \quad \rho(\tfrac{1}{2}\pi) = \rho^*.$$

According to (53₁)–(53₂), the function (52) has the constant (49) as minimum. Hence, while the expansion (51) was seen to be valid for every ζ whenever $0 < e < 1$ (and, actually, even in the limiting case $e = 1$ of periodic collisions), the expansion (50) cannot be used for all values of the time variable (7₂) unless the eccentricity e lies between 0 and $\rho^* = 0.6627 \dots$, a constant essentially less than 1. However, e is quite close to 0 in the majority of relevant astronomical applications.

§287. In order to prove (53₁), let σ denote any fixed positive number which is less than ρ^* . Then $|g(i\sigma)| < 1$, by (48). Hence, there exists, by (46₁), a positive $\theta < 1$ such that $|J_m(im\sigma)| < \text{const. } \theta^m$. Since (47₂) and (17₂) imply that $|J_m(mz)| < |J_m(im\sigma)|$ for $|z| < \sigma$, it follows that $|J_m(mz)| < \text{const. } \theta^m$ in the circle $|z| < \sigma$. Hence, it is clear from $0 < \theta < 1$ that if ζ has any fixed real value, the series

$$(54) \quad u \equiv u(z; \zeta) = \zeta + 2 \sum_{m=1}^{\infty} \frac{J_m(mz)}{m} \sin m\zeta$$

is uniformly convergent in the circle $|z| < \sigma$ of the complex z -plane. On the other hand, the functions $J_m(mz)$, where $m = 1, 2, \dots$, are regular analytic in the whole z -plane, since so are, by (17₂), the functions $J_m(z)$. Consequently, the series (54) represents, for every fixed

real ζ , a regular analytic function of z in the circle $|z| < \sigma$. Accordingly, (54) can be developed into a power series

$$(55) \quad u(z; \zeta) = \sum_{j=0}^{\infty} c_j(\zeta) z^j$$

which is valid, for every fixed real ζ , in the circle $|z| < \sigma$. Since σ was chosen as any positive number which is less than ρ^* , and since (54), (55) go over into (51), (50) by placing $z = e$, the proof of (53₁) is complete.

§288. There remains to be verified the relation (53₂), which will prove that ρ^* in (53₁) cannot be replaced by a number smaller than ρ^* , if all values of the angular time variable (7₂) are allowed.

First, if e is any fixed positive number, then either both series

$$(56_1) \quad \sum_{m=0}^{\infty} \frac{(-1)^m J_{2m+1}(i(2m+1)e)}{i(2m+1)};$$

$$(56_2) \quad \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{(m + \frac{1}{2})^{2m+2n+1} e^{2m+2n+1}}{n!(2m+n+1)!(2m+1)}$$

diverge to $+\infty$ or both converge to one and the same positive value. This is clear from the expansion (47₂), which is valid for every $|z|$ and implies that

$$(57) \quad i \cdot (-1)^m J_{2m+1}(i(2m+1)e) = |J_{2m+1}(i(2m+1)e)| > 0,$$

since $e > 0$; so that the terms of (56₁) and (56₂) are positive, and can, therefore, be arranged arbitrarily. For the same reason, (56₂) can be reordered into a simple power series

$$(58) \quad \sum_{n=0}^{\infty} a_n e^{2n+1}, \quad (a_n = \text{const.} > 0);$$

so that the three positive series (56₁), (56₂), (58) are, for a fixed $e > 0$, either all divergent to $+\infty$ or all convergent to one and the same number. Since (46₁), (48) and (57) show that the series (56₁) is convergent for $e < \rho^*$ and divergent for $e > \rho^*$, it follows that the same holds for the series (58), which is the Taylor series of the function (56₁). But this function (56₁) is, in view of (54), identical with the product of a constant ($= -\frac{1}{2}i$) and of $u(ei; \frac{1}{2}\pi) +$ another constant ($= -\frac{1}{2}\pi$). Consequently, the series (55) becomes at $\zeta = \frac{1}{2}\pi$

identical with (58), if one puts $z = ei$. Since the power series (58) in $e (> 0)$ was seen to be convergent or divergent according as $e < \rho^*$ or $e > \rho^*$, and since the radius of convergence of (55) at $\zeta = \frac{1}{2}\pi$ is $\rho(\frac{1}{2}\pi)$ by definition, the proof of (53₂) is complete.†

§289. The explicit form of the initial partial derivatives $c_j(\zeta)$ which are the coefficients of (50) can be obtained by using Lagrange's rule of differentiation. This rule states that

$$(59) \quad G(u) = G(\zeta) + \sum_{j=1}^{\infty} \frac{e^j}{j!} \frac{d^{j-1}}{d\zeta^{j-1}} \left\{ [H(\zeta)]^j \frac{d}{d\zeta} G(\zeta) \right\},$$

if the three variables u, ζ, e are subject to the relation

$$(60) \quad u = \zeta + eH(u).$$

Choosing, for instance, $G(v) = v$, one concludes from (59) that

$$(61) \quad u = \zeta + \sum_{j=1}^{\infty} \frac{e^j}{j!} \frac{d^{j-1}}{d\zeta^{j-1}} [H(\zeta)]^j$$

in virtue of (60). Needless to say, it is assumed that the given functions H, G are such that the expansion (59) is possible. For instance, (61) assumes that the given function $H(u)$ is such that (60) implicitly defines u , for a fixed ζ , as an analytic function of e , and that this analytic function has a regular analytic branch which becomes ζ at $e = 0$. Then this branch can, of course, be developed, for small $|e|$, into a Taylor series. Thus, Lagrange's rule (61) states merely that if ζ is fixed and $j = 1, 2, \dots$, then the j -th derivative with respect to e of the branch u under consideration becomes at $e = 0$ identical with the $(j-1)$ -th derivative of the j -th power of the given function $H(u)$ at $u = \zeta$; a fact which is easily verified by successive differentiations of the defining implicit relation (60).

§290. Let, in particular, $H(u) = \sin u$. Then (60) reduces to (7₃); hence, (61) to (50). Thus, comparison of (50) with (60), where $H(\zeta) = \sin \zeta$, gives $c_j(\zeta) = d^{j-1} \sin^j \zeta / d\zeta^{j-1}$ for $j = 1, 2, \dots$, while $c_0(\zeta) = \zeta$. But the j -th power of $\sin \zeta$ is, by de Moivre's rule, a

* The above proof (§287–§288) of (53₁)–(53₂) consisted in first establishing (53₁) and then (53₂). However, the coefficients of (58) are positive; so that it is clear from §288 that one could have established first (53₂) and then (53₁); and that (53₂) may be established directly if use is made of the function-theoretical fact that a power series $\sum a_n z^n$ which has a finite positive radius of convergence, r , and real non-negative coefficients a_n must represent a function which has a singularity at $z = r$ (Vivanti-Pringsheim).

linear combination of $1, \cos \zeta, \dots, \cos j\zeta$ or of $\sin \zeta, \dots, \sin j\zeta$ according as j is even or odd; so that the $(j-1)$ -st derivative of $\sin^j \zeta$ is a linear combination of $\sin \zeta, \dots, \sin j\zeta$ in both cases. On carrying out this calculation, one easily finds that

$$(62) \quad c_j(\zeta) = \frac{d^{j-1} \sin^j \zeta}{d\zeta^{j-1}} \equiv \sum_{k=0}^{[\frac{1}{2}j]} (j-2k)^{k-1} \frac{(-1)^k}{2^{j-1}} \binom{j}{k} \sin(j-2k)\zeta;$$

$$(j = 1, 2, \dots; c_0(\zeta) = \zeta),$$

$[\frac{1}{2}j]$ denoting the integral part of $\frac{1}{2}j$. Accordingly,

$$(63) \quad c_0(\zeta) = \zeta; \quad c_j(\zeta) = \sum_{l=1}^j \gamma_{jl} \sin l\zeta, \quad (j = 1, 2, \dots),$$

the γ_{jl} being numerical constants defined by (62).

This completes the explicit determination of the coefficients of the expansion (50) discussed in §285–§288.

§291. In §287–§288, the criterion (53₁)–(53₂) for the validity of the power series solution (50) of (7₃) was deduced from asymptotic properties of the coefficient functions of the Fourier series solution (51) of (7₃). Actually, one can arrive at (50) and (53₁)–(53₂), without following the detour via the Fourier series, if one applies to (7₃) the theory of analytic functions, as follows:

On placing

$$(64) \quad F(u, e; \zeta) = u - e \sin u - \zeta,$$

one can write (7₃) as $F(u, e; \zeta) = 0$; so that the problem is, for a fixed real ζ , the determination of that regular branch of the multi-valued analytic function $u = u(e; \zeta)$ defined by $F = 0$ for which $u(0; \zeta) = \zeta$; in fact, $F(u, 0; \zeta) = u - \zeta$, by (64). But the partial derivative of (64) with respect to the complex variable u is $F_u(u, e; \zeta) \equiv 1 - e \cos u$. Hence, $|F_u| > \text{const.} > 0$ as long as the complex variable u lies close enough to its real part and the complex variable e is sufficiently small in absolute value for any value of ζ . It follows, therefore, from the local existence theorem of analytic functions which are defined by an implicit condition $F = 0$, that $F(u, e; \zeta) = 0$, i.e. (7₃), defines the branch $u = u(e; \zeta)$ as a regular function element in e , with an expansion (50) which not only has a non-vanishing radius of convergence $\rho = \rho(\zeta)$ for every fixed real ζ but is, in addition, such that (53₁) holds for a sufficiently small positive constant ρ^* . That

(53₁) holds for the numerical constant (49) defined by (48), can be shown by an explicit discussion of the equations $F = 0$, $F_u = 0$ defined by (64). And the same direct discussion of the “nearest singularities” on the Riemann surface of $u = u(e; \zeta)$ which belongs to a fixed real ζ proves (53₂) also (cf., however, the end of §292).†

§292. It should be mentioned that the *implicit* problem $F = 0$ of §290–§291 reduces to an *inversion* problem. In fact, if one places

$$(66) \quad f(u; \zeta) = \frac{u - \zeta}{\sin u},$$

Kepler’s equation (7₃) appears in the form $e = f(u; \zeta)$. Hence, Kepler’s problem, i.e., the determination of $u = u(e; \zeta)$, simply is the problem of determining the inverse function of the meromorphic function (66) of u for every fixed real ζ . It is understood (cf. §291) that what matters in (50) and (53₁)–(53₂) is that branch of the inverse function $u = u(e; \zeta)$ of (66) for which $u(0; \zeta) = \zeta$. This proviso is necessary, since the meromorphic function (66) is transcendental, and so the Riemann surface of its inverse has, for every fixed ζ , infinitely many sheets. Correspondingly, the number $\rho(\zeta)$ occurring in (53₁) is the distance between $e = 0$ and the nearest singularity of $u = u(e; \zeta)$ on that sheet of this Riemann surface over the e -plane for which the numerator of (66) vanishes at $e = 0$.

The finite singularities of the inverse of a meromorphic function are known to be either algebraic branch points or transcendental singularities. The former depend on the zeros of the derivative, the

† The direct proof of (53₁)–(53₂) just indicated played an important historical rôle in the theory of analytic functions.

Lagrange derived his expansion (50), (62) of the solution of Kepler’s problem (7₃) only in a formal way, and did not prove the validity of (50), (62) even for $e < \frac{1}{10000}$, say. Several decades later, Laplace thought that he had succeeded in filling in this gap, and he arrived also at (53₁)–(53₂). Actually, the considerations of Laplace are purely heuristical and do not even prove that $\rho(\zeta) > \frac{1}{10000}$, say. This failure is quite understandable, since the problem is one which can be treated only by realizing the rôle played by the behavior of the functions in the complex domain (cf. the remarks on (ii) in §286); a point of view which was not at the disposal of Laplace. In fact, a principal impetus for Cauchy’s discoveries in complex function theory was his desire to find a satisfactory treatment for Lagrange’s series.

Cauchy was led to his fundamental theorem connecting the radius of convergence with the location of the nearest singularity, as well as to his maximum principle, precisely in his papers dealing with (53₁)–(53₂). Also the facts usually referred to as the argument principle and Rouché’s theorem were first observed in connection with this problem concerning Kepler’s equation.

latter on the asymptotic values, of the meromorphic function.* In the traditional proof of (53₁)–(53₂), which proceeds along the lines of §291 and is usually presented in text-books, only the zeros of the derivative of the function (66) of u are taken into account (for a fixed real ζ); so that the proof of (53₁)–(53₂) remains incomplete.†

§293. However, the omission can easily be corrected, since it turns out that asymptotic values do not matter in the present case. In fact, the entire function $\sin u$ of u does not have a (finite) asymptotic value. Hence, the meromorphic function (66) of u has 0 and only 0 as asymptotic value (for every fixed value of ζ). Consequently, the inverse function $u = u(e; \zeta)$ of $e = f(u; \zeta)$ cannot have a transcendental singularity at a finite e except at $e = 0$. But this transcendental singularity at $e = 0$ cannot belong to that sheet of the Riemann surface of $u = u(e; \zeta)$ in which one is interested, since on this sheet $u(e; \zeta)$ is regular at $e = 0$. Thus, the proof of (53₁)–(53₂) depends only on the determination of the algebraic singularities of $u = u(e; \zeta)$, it being understood that these singularities must be chosen on the relevant sheet.

§294. Consider again the method of §284–§289. Let the complex variable z be again restricted by $|z| < 1$, and let g be defined for $|z| < 1$ by (1₂); so that

$$(67) \quad \begin{aligned} g(z) &= \frac{z \exp (1 - z^2)^{\frac{1}{2}}}{1 + (1 - z^2)^{\frac{1}{2}}} \\ &\equiv \frac{|z| \exp \{ \psi i + (1 - |z|^2 \exp 2\psi i)^{\frac{1}{2}} \}}{1 + (1 - |z|^2 \exp 2\psi i)^{\frac{1}{2}}}, \end{aligned}$$

where $z = |z| \exp \psi i$, it being understood that $(1 - z^2)^{\frac{1}{2}} = +1$ at $z = 0$. Use will be made of the fact that‡ one has, besides (46₁),

$$(68) \quad |J_m(mz)| \leq |g(z)|^m.$$

A straightforward discussion of the elementary function (67) shows that the pair of conditions

* For instance, the inverse function of $w = \exp z$ has at $w = 0$ a logarithmic singularity which corresponds to the single asymptotic value $w = 0$ of $\exp z$.

† In view of the footnote to §291, it is worth mentioning that this omission in the usual proof of (53₁)–(53₂) was observed by Hurwitz when he introduced the theory of asymptotic values.

‡ This is shown in the theory of Bessel functions (Kapteyn series).

$$(69) \quad |g(R \exp \psi i)| = 1; \quad |g(|z| \exp \psi i)| < 1 \quad \text{if} \quad |z| < R$$

defines R as a unique continuous function of the angular variable ψ ; and that this $R = R(\psi)$ has the properties

$$(70) \quad R(\psi) = R(\psi + \pi) = R(-\psi), \quad (-\infty < \psi < +\infty);$$

finally, that

$$(71) \quad R(\psi_1) > R(\psi_2) \quad \text{if} \quad 0 < \psi_1 < \psi_2 \leq \frac{1}{2}\pi.$$

Now, (48) and (69) imply that $R(\frac{1}{2}\pi) = \rho^*$; while (3) shows that $R(\psi) \rightarrow 1$ as $\psi \rightarrow 0$. Hence, it is clear from (70) and (71) that if Γ denotes the curve $z = R(\psi) \exp \psi i$ in the complex plane $z = |z| \exp \psi i$, the region surrounded by Γ has the shape of a bi-symmetric convex

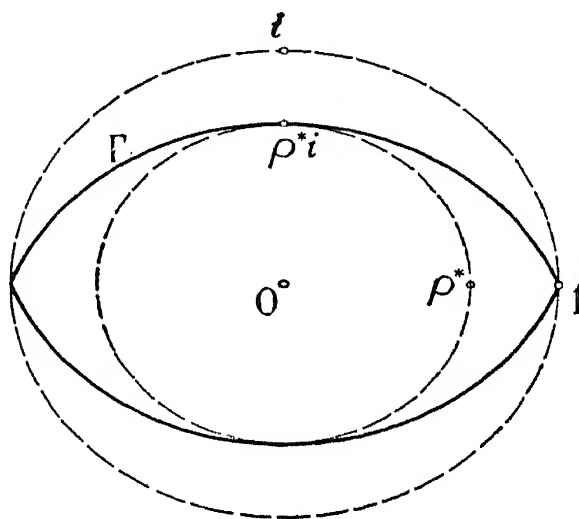


FIG. 9

lens which is contained in the circle $|z| < 1$, contains the circle $|z| < \rho^*$, and is, in view of (69), characterized by the pair of conditions

$$(72) \quad |g(z)| = 1 \quad \text{on} \quad \Gamma; \quad |g(z)| < 1 \quad \text{within} \quad \Gamma.$$

The corners of Γ on the real axis (cf. Fig. 9) are due to the algebraic branch points of (67) at $z = \pm 1$.

§295. It follows that the solution $u = u(z; \zeta)$ of Kepler's equation (7₃), where $z = e$ and $u(0; \zeta) = \zeta$, is regular analytic for every fixed real ζ not only in the circle $|z| < \rho^*$ (as proved in §287) but also in the larger domain which consists of the interior of the curve Γ (cf. Fig. 9). This is seen at once if, starting with (54), one repeats the considerations of §287 with the modification which consists of applying (68) and (72) instead of (46₁) and (48), respectively.

Since $u(z; \zeta)$ is for every real ζ regular in the z -domain consisting of the interior of Γ , it is seen from Fig. 9 that there exists for every positive $e_0 < 1$ a positive $\kappa = \kappa(e_0)$ such that $u(e; \zeta)$ can be developed according to powers of $e - e_0$ into a power series which has coefficients depending on ζ but is valid for every ζ as long as $|e - e_0| < \kappa(e_0)$. This result implies (53₁), as seen from Fig. 9 by letting $e_0 \rightarrow 0$.

§296. It follows also that there exists for $u = u(e; \zeta)$ an expansion which is valid for every $e = z$ within Γ and for every real ζ .

For let the function $Z = Z(z)$ map the interior of Γ upon the interior of the unit circle in the Z -plane in a one-to-one conformal manner. Then $u(z; \zeta)$ becomes, in virtue of the mapping, a function of $(Z; \zeta)$ which is regular analytic for $|Z| < 1$ and for every fixed ζ ; so that one has for $|Z| < 1$ a convergent Taylor expansion

$$(73) \quad u(z; \zeta) = \sum_{n=0}^{\infty} A_n Z^n, \text{ where } A_n = A_n(\zeta), Z = Z(z); (|Z| < 1).$$

Actually, one can choose $Z(z) = g(z)$. In fact, comparison of (70), (71) with (72), (69), where $R = R(\psi)$, shows that the curve Γ in the z -plane and the circle $|Z| = 1$ in the Z -plane are in one-to-one continuous correspondence if one puts $Z = g(z)$. Since the function (67) is regular analytic for $|z| < 1$ and so, by Fig. 9, in the interior of Γ , it follows from a standard lemma on conformal mapping (Darboux), that $Z = g(z)$ is a one-to-one conformal mapping of the interior of Γ upon $|Z| < 1$; q.e.d.

Now, the interior of Γ contains, by Fig. 9, the interval $0 \leq z = e < 1$. Hence, on placing $Z = g$ and $z = e$ in (73), one sees that the expansion

$$(74) \quad u(e; \zeta) = \sum_{n=0}^{\infty} A_n(\zeta) [g(e)]^n \quad [\text{cf. (1}_2\text{)}]$$

is, in the same way as (51) and in contrast to (50), valid for every positive $e < 1$ and for every real ζ .

§297. That the validity of (50) is, while that of (74) is not, restricted by the conditions represented by (53₁)–(53₂) and (49), can be explained by the fact that (50) and (74) are two different rearrangements of one and the same formal double series. This double series is obtained by developing the function (1₂), as well as its powers $[g(e)]^2, [g(e)]^3, \dots$, according to powers of e , and then rearrang-

ing (74) into (50) in a formal way. Using the explicit representation (62) of the coefficients of (50), one also obtains in this manner the explicit representation of the coefficients of (74).

§298. There is a similar explanation (besides the one given in §286) for the fact that the validity of (50) is, while that of (51) is not, restricted by the conditions represented by (53₁)–(53₂).

First, substitution of (62) or (63) into (50), when followed by a formal reordering, gives a double series of terms $\gamma_{jl}e^j \sin l\zeta$ (plus the single term $c_0 \equiv \zeta$), where the γ_{jl} are numerical constants. On the other hand, (17₂) shows that (51) can be written formally as a double series of the same form. Now, on applying (46₁), (47₂), (48) to the latter double series and otherwise proceeding in the same way as in §287–§288, one readily sees by a consideration of majorants, that the double series belonging to (51) is absolutely convergent, and can, therefore, be rearranged into (50), if $e(> 0)$ is less than ρ^* , while $\zeta(\geq 0)$ is arbitrary. Thus, the point† is that the rearrangement (51) is more favorable to convergence than the rearrangement (50), since (51) holds for every ζ if $0 \leq e < 1$, and not only if $0 \leq e < \rho^*$ ($= 0.662 \dots$).

§299. The object of §283–§298 was the investigation of the expansion of the Fourier series (22₁) according to the powers of the eccentricity. The behavior of the corresponding expansions of the remaining Fourier series of §278 is quite similar.

For instance, (5₁) shows that in order to develop the Cartesian coordinates according to powers of e into power series whose coefficients are functions of (7₂), it is sufficient to do the same for $\exp iu$. But the expansion (50) of u is valid on the assumptions expressed by (53₁)–(53₂); while $\cos u$ and $\sin u$ are entire functions of u and have zeros which clearly cannot compensate those singularities of (50) to which (53₂) is due. Hence, the expansion of $\exp iu$ in question, and so the corresponding expansion of the functions (5₁), is or is not valid for every value of the angular time variable (7₂) according as the integration constant $e(\geq 0)$ is less or greater than the number (49). Finally, the explicit form of the expansion in question is

† The formal identity of (50) and (51) was observed by Lagrange, who proceeded in reverse direction. In fact, Lagrange (cf. the footnote to §291) first found the power series (50) of restricted validity which he then formally re-ordered, with the help of (62), into the Fourier series (51); thus arriving at the transcendents (17₂) which to-day are called Bessel functions (cf. the end of §277).

$$(75) \quad \exp iu = \exp \zeta i + \sum_{j=1}^{\infty} \frac{e^j}{j!} \frac{d^{j-1}}{d\zeta^{j-1}} (\sin^j \zeta \exp \zeta i),$$

as seen by identifying (60) with (7₃) and placing $G(u) = \exp iu$ in (59).

Synodical Coordinates

§300. Let the coordinate system (x, y) of §258 be now denoted by (\bar{x}, \bar{y}) ; so that (1₂)–(2₂), §258 have to be written as

$$(1_1) \quad L = \frac{1}{2}(\bar{x}'^2 + \bar{y}'^2) + r^{-1};$$

$$(1_2) \quad \frac{1}{2}(\bar{x}'^2 + \bar{y}'^2) - r^{-1} = h; \quad (1_3) \quad \bar{x}\bar{y}' - \bar{y}\bar{x}' = c,$$

where $h \leq 0$, $c \geq 0$ and $r = (\bar{x}^2 + \bar{y}^2)^{\frac{1}{2}}$; while (15₁)–(15₃), §263 become

$$(2_1) \quad \bar{x} = r \cos (w + \omega), \quad \bar{y} = r \sin (w + \omega);$$

$$(2_2) \quad \omega = (w)_{t=t_0}; \quad (\min r(t) = (r)_{t=t_0}).$$

The Hamiltonian function belonging to (1₁) is seen to be

$$(3_1) \quad \bar{H} = \frac{1}{2}(\bar{X}^2 + \bar{Y}^2) - r^{-1}, \quad (r^2 = \bar{x}^2 + \bar{y}^2); \quad (3_2) \quad \bar{X} = \bar{x}', \quad \bar{Y} = \bar{y}'.$$

Introduce instead of the Cartesian coordinate system (\bar{x}, \bar{y}) another, (x, y) , which rotates about $(\bar{x}, \bar{y}) = (0, 0)$ with the constant angular velocity -1 ; so that

$$(4) \quad \bar{x} = x \cos t - y \sin t, \quad \bar{y} = x \sin t + y \cos t.$$

For reasons which will become apparent in §517, the rotating coordinate system (x, y) is called synodical, and the non-rotating (\bar{x}, \bar{y}) sidereal.

According to §95, the Lagrangian function in terms of (x, y) is

$$(5_1) \quad L = \frac{1}{2}(x'^2 + y'^2) + (xy' - yx') + (r^{-1} + \frac{1}{2}r^2); \quad (5_2) \quad r^2 = x^2 + y^2,$$

since (5₁) is readily seen to be identical with (1₁) in virtue of (4) and (5₂). In view of §229, the Hamiltonian function belonging to (5₁) is

$$(6_1) \quad H = \frac{1}{2}(X^2 + Y^2) - (xY - yX) - (r^{-1} - \frac{1}{2}r^2);$$

$$(6_2) \quad X = x' - y, \quad Y = y' + x.$$

Correspondingly, substitution of (4) into (1₂)–(1₃) gives

$$(7_1) \quad (x'^2 + y'^2) - (2r^{-1} + r^2) = -C; \quad (7_2) \quad -\frac{1}{2}C = h - c;$$

(cf. §210). It is clear from (6₁)–(6₂), and also from §155, that (7₁) is the energy integral of the irreversible dynamical system defined by (5₁), the energy constant with reference to the rotating coordinate system being denoted by $-\frac{1}{2}C$. This relative, or synodical, energy is, in view of (7₂), the difference of the sidereal energy h and of the (sidereal) angular momentum c .

§301. Consider, in particular, an arbitrary elliptic (incl. circular) path with the exclusion of segments; so that $h < 0$ and $c \neq 0$, by §242. Then, by (4), §241,

$$(8_1) \quad h = -\frac{1}{2}a^{-1}; \quad (8_2) \quad c^2 = a(1 - e^2),$$

while, by (15), §276,

$$(9_1) \quad n^2 = a^{-3}; \quad (9_2) \quad T = 2\pi:n,$$

where T is the (sidereal) period, and $c > 0$ or $c < 0$ according as the motion is direct or retrograde in the sidereal plane (§242). In view of (18), §265, this alternative may be expressed also by assigning the sign of n to be the same as that of c . Hence, (9₁)–(9₂) suggest the introduction of the square root $\alpha = \sqrt{a}$ with that determination for which α becomes of the same sign as n , i.e., as c . Thus, if A^\dagger denotes the positive square root for every $A > 0$, then

$$(10_1) \quad \alpha = a^\dagger \operatorname{sgn} c = \sqrt{a} \leq 0;$$

$$(10_2) \quad h = -\frac{1}{2}\alpha^{-2}; \quad (10_3) \quad c = \alpha(1 - e^2)^\dagger,$$

by (8₁), (8₂); while (9₁), (9₂) and (7₂) become

$$(11_1) \quad n = \alpha^{-3}; \quad (11_2) \quad T = 2\pi\alpha^3; \quad (11_3) \quad C = 2\alpha(1 - e^2)^\dagger + \alpha^{-2}.$$

Notice that the period (9₂) is defined to be of the same sign as c .

§302. Needless to say, the words “direct,” “retrograde” and “period” are meant in §301 in their sidereal sense, i.e., with reference to the non-rotating coordinate system (\bar{x}, \bar{y}) . The situation is quite different with reference to the synodical coordinate system (x, y) . Actually, an elliptic path, when considered from the rotating coordinate system, may be direct at some t and retrograde at some other t . In fact, substitution of (2₁)–(2₂) into (4) gives

$$(4 \text{ bis}) \quad x = r \cos(w - t + \omega), \quad y = r \sin(w - t + \omega), \quad (\omega = \text{const.}),$$

which shows that synodical orientation of a path at a given t is determined by the sign of the derivative $(w - t + \omega)'$, i.e., of the function

$w' = 1$ of t . Since (1₃) and (2₁)–(2₂) imply that $r^2 w' = c$, it follows that the motion is synodically direct or retrograde according as $c > r^2$ or $c < r^2$, where $c \neq 0$. But the maximum and the minimum of the focal radius vector $r = r(t)$ of an ellipse are $a(1 \pm e)$. Hence, the motion will pass from a synodically direct to a synodically retrograde orientation at a suitable $t = t^*$ if and only if the integration constants (8₁)–(8₂) are such as to make the constant (10₁) lie between the two positive bounds $a^2(1 \pm e)^2/(1 - e^2)^{\frac{1}{2}}$, where $0 < e < 1$; i.e., if and only if a is chosen between the two bounds $(1 \pm e)^{\frac{1}{2}}/(1 \mp e)$.

§303. Every sidereally retrograde ellipse is synodically retrograde for every t . This is clear from the criterion $c \geq r^2$ of §302, since $c > r^2$ cannot hold for $c < 0$.

On the other hand, a sidereally direct ellipse is synodically direct for every t only when a is less than $(1 - e)^{\frac{1}{2}}/(1 + e)$, where $0 \leq e < 1$, (which implies that $a < 1$). This follows by substituting into the condition $c > r^2$, where $c^2 = a(1 - e^2)$, the maximum of $r = r(t)$, which is $a(1 + e)$.

§304. Applying the criteria of §303 to the particular case $e = 0$, one sees that every sidereally retrograde circular path of radius a , where $0 < a < +\infty$, is synodically retrograde, and that in the sidereally direct case a circular path is synodically direct or retrograde according as $0 < a < 1$ or $1 < a < +\infty$. Finally, if the sidereally direct circular path is of radius $a = 1$, it is represented by a single point $x = \cos \omega$, $y = \sin \omega$ of the (synodical) circle $x^2 + y^2 = 1$, where ω is an arbitrary constant. In fact, if $\alpha \equiv \sqrt{a} = +1$, then $n = +1$, by (11₁); so that the sidereal circular motion has the constant angular velocity $+1$, and is, therefore, transformed by (4) to rest; cf. (4 bis), §302.

§305. The question of synodical periodicity will now be considered. In this respect, the case $0 < e < 1$ behaves quite differently from the circular case $e = 0$.

Excluding first the case $e = 0$ and noting that the t -period of the rotation (4) is 2π , one sees that the synodical path $x = x(t)$, $y = y(t)$ does or does not close into itself after the lapse of a sufficiently high number of sidereal periods (9₂), according as the value of the integration constant n is rational or irrational.

If n is irrational, the synodical path, where $-\infty < t < +\infty$, is everywhere dense on the circular ring having the radii $\max r(t) = a(1 + e)$ and $\min r(t) = a(1 - e)$, the reason being the same as in §215 (where the circular ring becomes a circular disk, illustrated in

the figure below). If, on the other hand, n is rational, say $n = p:q$, where p, q are integers, then the synodical path closes into itself after

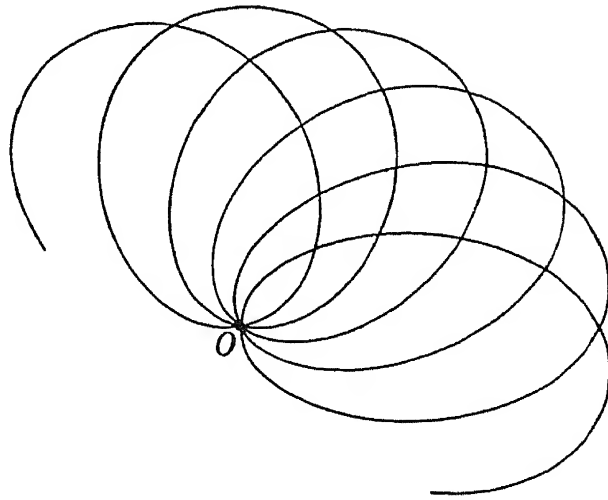


FIG. 10

the lapse of $|p|$ sidereal periods (and, if p, q are relatively prime, not earlier). In fact, it is clear from (4) and (9₂) that if τ denotes the primitive synodical period, then

$$(12) \quad \pm \tau = pT = 2\pi q, \text{ where } n = p:q, (p, q) = 1; \quad (e \neq 0).$$

In particular, the primitive sidereal and synodical periods, T and τ , are of equal magnitude only when the period 2π of the rotation (4) divides T , i.e., when and only when n is the reciprocal value of some integer q .

§306. It will be shown in §307 that the situation is quite different in the circular case $e = 0$, since in this case the synodical circular motion is periodic and has the primitive period

$$(13) \quad \tau^* = 2\pi/(n - 1), \quad (e = 0),$$

whether n is irrational or rational.

Notice that (13) differs from (12) when (12) exists, i.e., when n is rational. In fact, (13) then reduces to

$$(13 \text{ bis}) \quad \tau^* = 2\pi q/(p - q), \text{ where } n = p:q, (p, q) = 1; \quad (e = 0).$$

Thus, if the value of (11₁) is fixed and is rational, and if e varies, then (12) is, for all $e \neq 0$, independent of e and, therefore, identical with $\lim \tau$ as $e \rightarrow 0$; while (12) and (13 bis) show that this circular limit, $\lim \tau$, of the non-circular primitive synodical period, instead of being the circular primitive synodical period τ^* , is a multiple of τ^* ;

namely, $(p - q)\tau^*$. (This discontinuity becomes important in the theory of the periodic solution of the restricted problem of three bodies).

Of particular interest are those accidental values of $n = p:q$ for which this discontinuity does not arise, i.e., for which $p - q = +1$. The corresponding values of $n = \alpha^{-3} \equiv \sqrt{a^{-3}}$, namely, the values $n = p:(p + 1)$, where $p = 1, 2, \dots$ and $p = -2, -3, \dots$, will be referred to as critical. Notice that the assumption $p - q = -1$, under which $\lim \tau = -\tau^*$, leads to the same critical values.

It is tacitly assumed that $n \neq 1$, since (13) becomes meaningless for $n = 1$. Actually, $n = 1$ is the case mentioned at the end of §304; so that in this case the synodical circular period is arbitrary, since the synodical circular motion becomes an equilibrium solution.

§307. In order to prove (13) for rational and irrational n , notice that the circular sidereal motions $\bar{x} = \bar{x}(t)$, $\bar{y} = \bar{y}(t)$ are uniform rotations, with n as angular velocity; so that $\bar{x} = a \cos nt$, $\bar{y} = a \sin nt$ upon a suitable choice of the origin of the t -axis. Thus, (4) shows that the synodical path is $x = a \cos (n - 1)t$, $y = a \sin (n - 1)t$. This implies (13), and also the exceptional behavior for $n = 1$.

It will be convenient to write (13) as $\tau^* = 2\pi m$; so that

$$(14_1) \quad m = \frac{1}{n - 1}; \quad (14_2) \quad \alpha \equiv \sqrt{a} = \frac{m^{\frac{1}{3}}}{(1 + m)^{\frac{1}{3}}}; \quad (14_3) \quad C = \frac{1 + 3m}{m^{\frac{1}{3}}(1 + m^2)^{\frac{1}{3}}},$$

by (11₁) and (11₃), where $e = 0$ and $n \neq 1$. In the exceptional case, (11₁) and (11₃) show that

$$(15) \quad \alpha \equiv \sqrt{a} = 1 \quad \text{and} \quad C = 3 \quad \text{for} \quad n = 1; \quad (e = 0).$$

Notice that m is rational if and only if so is n , and that m is an integer ($\neq 0$) if and only if n is critical in the sense of §306. This is clear from the definition (14₁) of the continuous parameter $m \geq 0$.

§308. If the value of the sidereal energy constant $h (< 0)$ is given, there exist exactly two circular paths corresponding to it. In fact, the square of $\alpha \equiv \sqrt{a} \geq 0$, which is the radius, then follows uniquely from (10₂); while the sidereal period is determined by (11₁)–(11₂). The situation is more involved if the full range $-\infty < \alpha < +\infty$, ($\alpha \neq 0$), of circular paths is to be described not in terms of the sidereal energy constant h but in terms of the synodical energy constant, which is $-\frac{1}{2}C$, by (7₁)–(7₂). In fact, this description, instead of being a simple one-to-two correspondence, is as follows:

Exclude the meaningless case $\alpha = 0$ of a vanishing radius $a = \alpha^2$, and the exceptional case (15) also. The pair $\alpha = 0, \alpha = 1$ of omitted values separates the full range $-\infty < \alpha < +\infty, (\alpha \neq 0)$, of circular paths into the three ranges

$$(16_1) \quad -\infty < \alpha < 0; \quad (16_2) \quad 0 < \alpha < 1; \quad (16_3) \quad 1 < \alpha < +\infty,$$

(16₁) and (16₂)–(16₃) representing the sidereally retrograde and direct circular paths, respectively. Now, the corresponding C -ranges are

$$(17_1) \quad -\infty < C < +\infty; \quad (17_2) \quad +\infty > C > 3; \quad (17_3) \quad 3 < C < +\infty,$$

with the understanding that the correspondence between the ranges (16 _{ν}) and (17 _{ν}) is one-to-one for every ν ($= 1, 2, 3$), and that the manner of writing in (17 _{ν}) indicates also the increase or decrease of the function $C = C(\alpha)$ on the range (16 _{ν}). For instance, (16₂), (16₃) and (17₂), (17₃) show that C tends to 3, whether α tends decreasingly or increasingly to the exceptional value (15) of α ; while (16₁), (17₁) imply that there belongs to $C = 3$ a non-exceptional $\alpha (< 0)$ also. Incidentally, the latter is $\alpha = -\frac{1}{2}$, since elimination of m between (14₂)–(14₃) gives $C = 2\alpha + \alpha^{-2}$ for any α ; (cf. (11₃) for $e = 0$).

Now, the derivative of the function $C = 2\alpha + \alpha^{-2}$ of α is $2(1 - \alpha^{-3})$. Hence, $C = C(\alpha)$ is strictly increasing or strictly decreasing on the α -range (16 _{ν}) according as $\nu = 1, 3$ or $\nu = 2$. Since $C(\alpha) = 2\alpha + \alpha^{-2}$ implies also that $C(\pm\infty) = \pm\infty, C(\pm 0) = +\infty, C(1 \pm 0) = 3$, the proof of (16₁)–(17₃) is complete.

§309. Since $\alpha = 0$ is excluded, one can write $C = 2\alpha + \alpha^{-2}$ as a cubic equation either for α or for $1/\alpha$, as follows:

$$(18_1) \quad 2\alpha^3 - C\alpha^2 + 1 = 0; \quad (18_2) \quad (1/\alpha)^3 - C \cdot 1/\alpha + 2 = 0.$$

And (16₁)–(17₃) imply that the cubic equation (18₁) has

- (i) exactly one negative root, say $\alpha = \alpha_-(C)$, for $-\infty < C < +\infty$;
- (ii) no positive root α if $C < 3$, and two distinct positive roots, say $\alpha_+ = \alpha_+(C)$ and $\alpha^+ = \alpha^+(C)$ for $C > 3$, where $\alpha_+ < 1 < \alpha^+$ for $C > 3$, and $\alpha_+ \rightarrow 1 - 0, \alpha^+ \rightarrow 1 + 0$ as $C \rightarrow 3 + 0$.

Actually, the discriminant of the cubic equation (18₂) is seen to be $-4(-C)^3 - 27 \cdot 2^2$; this may be written as $4(C^3 - 3^3)$ and is, therefore, ≥ 0 according as $C \geq 3$. Thus, (i)–(ii) follow not only from (16₁)–(17₃) but also directly from (18₂) or (18₁).

It should be mentioned for later reference that

$$(19) \quad \alpha_-^2 < \alpha_+^2; \quad \alpha_-^2 > 1/\alpha_+, \quad \alpha_+^2 < 1/\alpha^+, \\ (C > 3; \quad \alpha_- < 0 < \alpha_+ < 1 < \alpha^+).$$

In fact, on using (16₁)–(17₃) and the definitions (i)–(ii) of α_{\pm}^2 for $C > 3$, one readily verifies* the three inequalities (19) either directly or by differentiation of (18₁)–(18₂).

§310. The limiting case $c = 0$ (i.e., $e = 1$) of a sidereal elliptic motion of arbitrary major axis $2a$ will now be considered.

According to §268, the parametrization of these rectilinear sidereal motions in terms of the eccentric anomaly u may be written as

$$(20_1) \quad \bar{x} = a(\cos u - 1), \quad \bar{y} \equiv 0; \quad (20_2) \quad t = (u - \sin u)/n; \\ (20_3) \quad n^2 a^3 = 1,$$

with the understanding that the sign of n , i.e., the sidereal orientation, cannot be defined. From (8₁)–(13₃),

$$(21_1) \quad T^2 a^{-3} = 4\pi^2; \quad (21_2) \quad 1/a = -2h = C > 0,$$

since $c = 0$. According to (20₁)–(21₁), the t -period T is the amount of time elapsing between two successive collisions of the moving particle with the body which rests at $(\bar{x}, \bar{y}) = (0, 0)$; cf. §268–§270.

Substitution of (20₁)–(20₃) into (4) shows that the sidereal path is

$$(22) \quad x = -2a \sin^2 \frac{1}{2}u \cos (a^{\frac{3}{2}}u - a^{\frac{3}{2}} \sin u), \\ y = 2a \sin^2 \frac{1}{2}u \sin (a^{\frac{3}{2}}u - a^{\frac{3}{2}} \sin u),$$

where the auxiliary time variable u runs from $-\infty$ to $+\infty$. According to (22), the collisions (i.e., the states with $x^2 + y^2 = 0$) occur at the equidistant u -dates $u = 0, \pm 2\pi, \dots$. Nevertheless, the synodical path is not, in general, a closed curve. In fact, (22) shows that the full (x, y) -path will or will not be a closed curve (having a sufficiently high number of “loops” or “circuits”) according as the value of the integration constant $n = a^{-\frac{3}{2}}$ is rational or irrational. In the first case, (20₃), (21₁) and (22) imply that the relation (12), originally derived for $0 < e < 1$, holds also in the present case $e = 1$ of a rotating segment (of length $2a$). And the result of §305 remains valid for the second case also, since the (x, y) -locus (22), where

* For instance, the first of the inequalities (19) is certainly true for those $C > 3$ which are very close to $C = 3$; in fact, $\alpha_-(3) = -\frac{1}{2}$, $\alpha_+(3) = 1$, by (18₁). Hence, if $\alpha_-^2 < \alpha_+^2$ were not true for all $C > 3$, there would exist, for reasons of continuity, a certain $C = C_0$ at which $\alpha_-^2 = \alpha_+^2$. But for such a $C = C_0$ one would have $\alpha_-^3 = \alpha_+^3$, by (18₁). And $\alpha_-^3 = \alpha_+^3$ is impossible, since $\alpha_- < 0 < \alpha_+$.

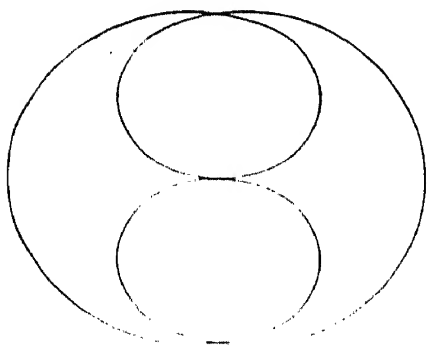
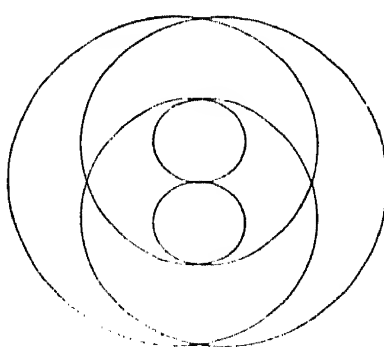
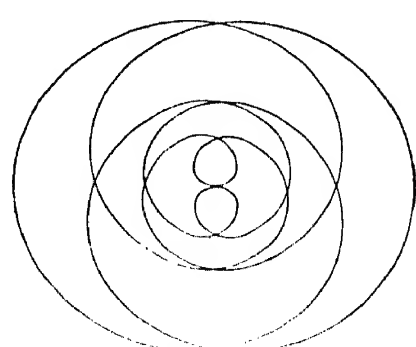
$-\infty < u < +\infty$, clearly is dense on the circle $x^2 + y^2 \leq (2a)^2$, if $n = a^{-\frac{1}{3}}$ is irrational.

§311. The fact that the synodical path is not periodic in case of an irrational n does not contradict the fact that the synodical path passes through the same point $(x, y) = (0, 0)$ infinitely often (namely, when $u = 0, \pm 2\pi, \dots$). The situation becomes quite intuitive by introducing synodical polar coordinates r, ϑ . For then (22) may be written as

$$(23) \quad \begin{aligned} x &= r \cos \vartheta, \quad y = r \sin \vartheta, \quad \text{where} \\ r &= 2a \sin^2 \frac{1}{2}u, \quad \vartheta = -a^{\frac{1}{3}}(u - \sin u) + \pi. \end{aligned}$$

If the time parameter u is increased by some multiple of 2π , say by $2\pi p$, then (23) shows that, while r remains unchanged, ϑ decreases by $2\pi p a^{\frac{1}{3}}$. And (20₃), (21₁) show that this decrease of the synodical polar angle ϑ is a multiple of 2π only when the commensurability condition (12) for n is satisfied by some integer q . Consequently, there will or will not exist among the dates $u = 0, \pm 2\pi, \dots$ a date at which the synodical path will leave the origin $(x, y) = (0, 0)$ in the same direction (mod 2π) as that in which it arrived at the same date, according as n is rational or irrational.

If n is rational, say $n = p:q$, where $(p, q) = 1$, then only certain, and not all, of the collision dates $u = 0, \pm 2\pi, \dots$ must be such that the angle ϑ remains unchanged (mod 2π) during the passage through the origin. In fact, ϑ will remain unchanged (mod 2π) during

FIG. 11₁FIG. 11₂FIG. 11₃

each of these passages only when the change in ϑ which corresponds to an increase of u by 2π happens to be a multiple of 2π , say $2\pi q$. In view of the representation (23) of ϑ , this will be the case if and only if $2\pi a^{\frac{1}{3}} = 2\pi q$. And this means, by (20₃), that $1/n = q$.

Accordingly, the entering and departing branches of the synodical path (22) touch each other at all dates of collision if and only if n is the reciprocal value of an integer q . These n belong, by (20₃), to the discrete values

$$(24) \quad C^{-1} \equiv a = q^{\frac{2}{3}}; \quad (q = 1, 2, \dots),$$

of the arbitrary integration constant (21₂), which always determines a bounded collision path uniquely. The synodical paths (22) which belong to $q = 1$, $q = 2$ and $q = 3$ are shown in the figures.*

§311 bis. Only the elliptic case $h < 0$ was considered in §301–§311. On substituting into (4) the sidereal coordinates \bar{x} , \bar{y} of an hyperbolic or a parabolic motion, one sees how the synodical (x, y) -path will behave if $h > 0$ or $h = 0$, where the limiting case of the respective collision paths ($c = 0$) is not excluded.

§312. If the value of the integration constant of the sidereal energy h is given, (1₂) does not or does determine a curve of zero velocity according as $h \geq 0$ or $h < 0$, and the configuration domain precluded in the latter case is the exterior of the circle of radius $-h^{-1}$ about $(\bar{x}, \bar{y}) = (0, 0)$; cf. §243. The corresponding discussion of the case in which (1₂) is replaced by its synodical analogue (7₁) is somewhat more involved, and proceeds as follows:

It is clear from (7₁) that, at every point (x, y) of any synodical solution path $x = x(t)$, $y = y(t)$ of given synodical energy (7₂), one must have $r^2 + 2r^{-1} \geq C$, and that $r^2 + 2r^{-1} = C$ is the equation of the corresponding curve of (synodical) zero velocity; an equation which represents as many circles about the origin $(x, y) = (0, 0)$ as is the number of its distinct positive roots r . But the equation $r^2 + 2r^{-1} = C$ appears in the form (18₁), if one puts $\alpha = 1/r$. And it was shown in §309 that (18₁) has no positive root, the single positive double root $\alpha = 1$ or exactly two positive roots $\alpha_+ = \alpha_+(C)$, $\alpha^+ = \alpha^+(C)$, where $\alpha_+ < 1 < \alpha^+$, according as $-\infty < C < 3$, $C = 3$ or $3 < C < +\infty$.

Consequently, the curve of synodical zero velocity, belonging to a given value of C , does not exist or consists of two concentric circles about the origin according as $C < 3$ or $C > 3$. Furthermore, the radii of the circles in the case $C > 3$ are $1/\alpha^+$ and $1/\alpha_+$, where $1/\alpha^+ < 1 < 1/\alpha_+$; and these two circles, which disappear for $C < 3$, coincide with the unit circle for $C = 3$.

* It is clear from (24) that the three figures are drawn on different scales of the unit of length.

In addition, the (x, y) -region prohibited by (7₁), i.e., the (x, y) -region in which $r^2 + 2r^{-1} \geq C$ does not hold, consists for $C > 3$ of the ring $1/\alpha^+ < (x^2 + y^2)^{\frac{1}{2}} < 1/\alpha_+$, which degenerates for $C = 3$ into the circle $x^2 + y^2 = 1$ (and disappears for $C \leq 3$). This becomes clear by observing that if C is arbitrarily fixed, the requirement $r^2 + 2r^{-1} \geq C$ of (7₁) is satisfied when $r > 0$ is very close either to $r = 0$ or to $r = \infty$; so that, α_+ and α^+ being (for $C > 3$) simple roots of the cubic equation (18₁), the condition $r^2 + 2r^{-1} \geq C$ cannot be satisfied in the ring $1/\alpha^+ < r < 1/\alpha_+$.

§312 bis. On comparing the results of §308–§309, which concern circular paths, with the results of §312, which concern the curves of zero velocity for arbitrary paths $x = x(t)$, $y = y(t)$, one sees from (19) the relative location of any circular path and of the ring precluded by its energy integral (if $C > 3$). It is also seen how the limiting case $C = 3$ of §309 corresponds to the exceptional case $n = 1$ of §306.

Let a circular path of radius a be called lower or upper according as $a < 1$ or $a > 1$. In either case, the path may be sidereally direct or retrograde ($\alpha \equiv \sqrt{a} \gtrless 0$). Thus, exclusion of $a = 0$ and of $a = 1$ cuts the full range of circular solutions into the four ranges

$$A_1: \quad 0 < \alpha < 1;$$

$$A_2: \quad 1 < \alpha < +\infty;$$

$$A_3: \quad -\infty < \alpha < -1;$$

$$A_4: \quad -1 < \alpha < 0$$

of $\alpha \equiv \sqrt{a} \gtrless 0$. The notations and results of §306–§309 concerning circular paths, together with that interpretation of (19) which follows from §312, may be collected as follows:*

* It is quite an accident that the two C -values under XV are nearly equal; they are by no means identical (as sometimes stated in the literature).

		A_1	A_2	A_3	A_4
I II III IV	$\alpha = \sqrt{a}$ α_- ; α_+ or α^+ $C = 2\alpha + \alpha^{-2}$ radius of circle of zero vel. (synod.)	$0 < \alpha < 1$ $\alpha = \alpha_+$ $+\infty > C > 3$ $0 < 1/\alpha^+ < 1$	$1 < \alpha < +\infty$ $\alpha = \alpha^+$ $3 < C < +\infty$ $1 < 1/\alpha_+ < +\infty$	$-\infty < \alpha < -1$ $\alpha = \alpha_-$ $-\infty < C < -1$ non-existent	$-1 < \alpha < 0$ $\alpha = \alpha_-$ $-1 < C < +\infty$, if $1 > 1/\alpha^+ > 0$, if $3 < C < +\infty$; non-existent, if $-\infty < C < 3$ $1/\alpha^+ > a$, if $3 < C < +\infty$; non-existent, if $-\infty < C < 3$ $a = \alpha_-^2$
V	location of circle of zero vel. (synod.)	$1/\alpha^+ > a$	$1/\alpha_+ < a$	non-existent	
VI	radius of circular orbit	$a = \alpha_+^2 (> \alpha_-^2)$	$a = \alpha^{+2} (> \alpha_-^2)$	$a = \alpha_-^2$	
VII VIII IX X XI	upper or lower orbit sidereal orient. synodical orient. $n = \alpha^{-3}$ $m = (n-1)^{-1}$	lower; $0 < a < 1$ direct direct $+\infty > n > 1$ $0 < m < +\infty$	upper; $1 < a < +\infty$ direct retrograde $1 > n > 0$ $-\infty < m < -1$	upper; $+\infty > a > 1$ retrograde retrograde $0 > n > -1$ $-1 < m < -\frac{1}{2}$	lower; $1 > a > 0$ retrograde retrograde $-1 > n > -\infty$ $-\frac{1}{2} < m < 0$
XII XIII XIV	critical m critical n cluster value of crit. a	$1, 2, 3, \dots$ $\frac{1}{2}, \frac{3}{2}, \frac{4}{3}, \dots$ $a = 1 - 0$	$\dots, -4, -3, -2$ $\dots, \frac{3}{4}, \frac{2}{3}, \frac{1}{2}$ $a = 1 + 0$	non-existent non-existent non-existent	non-existent non-existent non-existent
XV	first crit. C	$(C)_{m=1} = \sqrt[3]{32}$ $= 3.1748 \dots$	$(C)_{m=-2} = \sqrt[3]{31\frac{1}{4}}$ $= 3.1498 \dots$	non-existent	non-existent
XVI	beginning	$m = +0$, $n = +\infty, C = +\infty$	$m = -1 - 0$, $n = +0, C = +\infty$	$m = -1 + 0$, $n = -0, C = -\infty$	$m = -0$, $n = -\infty, C = +\infty$

CHAPTER V

THE PROBLEM OF SEVERAL BODIES

Newton's law of gravitation	§313–§321
Consequences of the conservation integrals	§322–§332
Simultaneous collisions	§333–§339
Heliocentric coordinates	§340–§347
Binary collisions	§348–§354
Central configurations	§355–§368
Homographic solutions	§369–§374
Homographic solutions and central configurations	§375–§382
Elimination of the linear momentum	§383–§389
Elimination of the angular momentum	§390–§406
Real singularities	§407–§414
The function-theoretical character of the collisions	§415–§425
The problem of three bodies	§426–§440

Newton's Law of Gravitation

§313. By saying that a system of $n(\geq 2)$ particles P_1, \dots, P_n is moving according to Newton's law of gravitation, one means that there exist

- (i) positive constants $\kappa; m_1, \dots, m_n$;
 - (ii) a suitably chosen Cartesian coordinate system $\xi = (\xi^I, \xi^{II}, \xi^{III})$ in a Euclidean 3-space;
 - (iii) a suitably chosen independent variable t ,
- such that the system of equations of motion can be written as

$$(1) \quad m_i \frac{d^2 \xi_i}{dt^2} = \left\{ \kappa \sum_{1 \leq j < k \leq n} \frac{m_j m_k}{|\xi_j - \xi_k|} \right\}_{\xi_i}, \quad (i = 1, \dots, n),$$

where ξ_i denotes the 3-vector of the coordinates $\xi_i^I, \xi_i^{II}, \xi_i^{III}$ of P_i , and $\{ \ }_{\xi_i}$ the ξ_i -gradient of the scalar $\{ \}$.

The parameters κ and m_i are, respectively, the "constant of gravitation" and the "mass of P_i "; while a coordinate system ξ and an independent variable t for which (1) is valid are termed an "inertial coordinate system" and an "absolute time" of Newton's theory. Needless to say, it is impossible to speak of any of these notions without involving each of the other notions. For instance, it is meaningless to ask what are the values m_i of the masses if one does not grant as known an inertial coordinate system and an absolute time.

It does not lie within the province of this book, to discuss the long series of incomparable triumphs and the few, though not negligible, failures of the above-described pre-Einsteinian approach to the problem of gravitation in the solar system. Hence, it will not be necessary to discuss the practical and logical difficulties which are involved by the (necessarily implicit) definition of an inertial coordinate system, when the mathematical model is applied to the motion of the planets and their satellites. The practical difficulty just mentioned is, of course, a purely astronomical problem. And the astronomical technique of the numerical embedding of all direct and indirect observational data into the Newtonian model is so highly developed, that the difficulty in question has no practical significance for the present state of the theory of the solar system.

In what follows, m_i will denote not only the mass concentrated at the variable point ξ_i but also the particle P_i whose coordinate vector is ξ_i . Similarly, ξ will denote not only the coordinate system but also the coordinate vector of a point of the 3-dimensional Euclidean space ξ . Finally, $\xi^I, \xi^{II}, \xi^{III}$ will denote the components of ξ parallel to the coordinate axes.

§314. Denoting by $u \times v = -v \times u$ the vector product of two 3-vectors u, v and by $u \cdot v = v \cdot u$ their scalar product, finally by u^2 the square $u \cdot u$ of the length $|u|$ of u , put

$$(2_1) \quad C = \sum m_i \xi_i \times \xi'_i; \quad (2_2) \quad J = \sum m_i \xi_i^2; \quad (2_3) \quad L = T + U,$$

where $' = d/dt$, the scalars T, U are defined by

$$(3_1) \quad T = \frac{1}{2} \sum m_i \xi_i'^2; \quad (3_2) \quad U = \sum^* m_j m_k / \rho_{jk}; \quad (3_3) \quad \rho_{jk} = |\xi_j - \xi_k|,$$

and the summation signs \sum, \sum^* by

$$(4_1) \quad \sum = \sum_{i=1}^n; \quad (4_2) \quad \sum^* = \sum_{1 \leq j < k \leq n}.$$

The 3-vector (2_1) is called the angular momentum, and the scalar (2_2) the polar inertia momentum. Notice that the masses m_i are constant scalars and that ξ_i^2 is the square of the distance $|\xi_i|$ between m_i and the origin $\xi = 0$, while (3_3) is the distance between m_j and m_k . The coordinate vector of the centre of mass is $\mu^{-1} \sum m_i \xi_i$, if $\mu = \sum m_i$ denotes the total mass.

Choosing the constant of gravitation κ to be unity, one sees from (2_3) – (4_2) that (1) can be written in the form

$$(5) \quad m_i \xi_i'' = U_{\xi_i} \equiv U_{\xi_i}(\xi_1, \dots, \xi_n) \text{ or } [L]_{\xi_i} = 0, \quad (i = 1, \dots, n)$$

$[]_{\xi_i}$ denoting Lagrangian differentiation with respect to the 3-vector coordinate ξ_i . Thus, (1) is a conservative dynamical problem which has $3n$ degrees of freedom and is, by (2₃)–(3₂), reversible (§156). The condition (2₁), §155 is satisfied, since $T = \frac{1}{2} \sum m_i \xi_i'^2$ is a diagonal form with positive diagonal elements $m_1, m_1, m_1; \dots; m_n, m_n, m_n$. Correspondingly, from (10)–(11₃), §158,

$$(6_1) \quad L_{\xi_i'} \equiv \eta_i = m_i \xi_i'; \quad (6_2) \quad H = \frac{1}{2} \sum m_i^{-1} \eta_i^2 - U; \quad (6_3) \quad T = \frac{1}{2} \sum m_i^{-1} \eta_i^2,$$

if η_i denotes the 3-vector whose components are the momenta canonically conjugate to the components of the coordinate 3-vector ξ_i , where $i = 1, \dots, n$.

§315. Similarly, the expression at the left of (14), §159 reduces to $(\sum m_i \xi_i \cdot \xi_i')'$, and so, by (2₂), §314, to $\frac{1}{2} J''$. On the other hand, the expression at the right of (14), §159 reduces to (15₁), §159, since $T = \frac{1}{2} \sum m_i \xi_i'^2$ is independent of, hence homogeneous of degree $\alpha = 0$ in, the coordinates. Finally, (3₂)–(3₃), §314 show that U is homogeneous of degree -1 ; so that (15₁), §159 reduces to (15₂), §159, with $\beta = -1$. Accordingly, $\frac{1}{2} J'' = (-1 + 2)U + 2h$, i.e.,

$$(7_1) \quad J'' = 2U + 4h; \quad (7_2) \quad T - U = h,$$

(7₂) being the definition of h in (7₁), i.e., of the energy constant of a given solution

$$(8) \quad \xi_i = \xi_i(t) \quad (i = 1, \dots, n).$$

§315 bis. According to §160, another consequence of $\beta = -1$ is that $\xi_i = \lambda \xi_i(\lambda^{-\frac{2}{3}} t)$ is, for every constant $\lambda > 0$ and for every solution (8) of (5), again a solution of (5). This implies, in particular (cf. §160 bis), that the period within a family of periodic solutions of (5) is proportional to $|h|^{-\frac{3}{2}}$, if the particular solutions which constitute the family have continuous partial derivatives with respect to its parameters.

§316. Replace the coordinate system ξ of §313 by another coordinate system, $\bar{\xi}$, which is obtained from ξ by a fixed rotation about the origin; so that $\bar{\xi}_i = \Omega \xi_i$, where Ω is an orthogonal 3-matrix which has the determinant $+1$ and is independent of t and i . Clearly, $\bar{\xi}_i'^2 = \xi_i'^2$ and $|\bar{\xi}_j - \bar{\xi}_k| = |\xi_j - \xi_k|$. Hence, (3₁)–(3₃) show that (2₃) is invariant under the transformation $\bar{\xi}_i = \Omega \xi_i$.

Consequently, if one replaces Ω by $\Omega(\epsilon)$, where ϵ is a scalar parameter independent of t and the orthogonal matrix function $\Omega(\epsilon)$ has at $\epsilon = 0$ a non-vanishing derivative $\Omega_\epsilon(0)$, while $\Omega(0)$ is the unit matrix, then $\sum L_{\xi'_i} \cdot \Omega_\epsilon(0) \xi_i$ is, according to (8), §96, an integral of (5), §314.

Choose, in particular, the family $\Omega = \Omega(\epsilon)$ of conservative rotations by placing $d_\nu = \epsilon \delta_\nu$ in (18₂), §77, where the scalars $\delta_1, \delta_2, \delta_3$ are arbitrarily fixed. Then the derivative $\Omega_\epsilon(0)$ is seen to be the sum of the three matrices $\delta_\nu I_\nu$. Substituting this sum into the integral $\sum L_{\xi'_i} \cdot \Omega_\epsilon(0) \xi_i$ and choosing successively $(\delta_1, \delta_2, \delta_3) = (1, 0, 0); (0, 1, 0); (0, 0, 1)$, one obtains the three integrals $\sum L_{\xi'_i} \cdot I_\nu \xi_i$, where $\nu = 1, 2, 3$. But the three skew-symmetric matrices I_ν , defined at the beginning of §77, are easily verified to have the property that if \mathbf{A}, \mathbf{B} are two 3-vectors, the three scalar products $\mathbf{B} \cdot I_\nu \mathbf{A}$ are the components of the vector product $\mathbf{A} \times \mathbf{B}$. Hence, the three components of the vector $\sum \xi_i \times L_{\xi'_i}$ are integrals of (5). Substituting $L_{\xi'_i}$ from (6₁), one can say that there exists for every solution (8) of (5) a constant vector C such that

$$(9) \quad \sum m_i \xi_i \times \xi'_i = C.$$

In other words, the angular momentum vector (2₁) represents three scalar integrals of (5).

§317. Replace the coordinate system ξ of §313 by another coordinate system, $\bar{\xi}$, which is obtained from ξ by a fixed translation; so that $\bar{\xi}_i = \xi_i + b$, where b is a 3-vector which is independent of t and i . Clearly, $\bar{\xi}'^2 = \xi'^2$ and $|\bar{\xi}_j - \bar{\xi}_k| = |\xi_j - \xi_k|$. Hence, (3₁)–(3₂) show that (2₃) is invariant under the transformation $\bar{\xi}_i = \xi_i + b$.

Replacing b by ϵc , where ϵ is a scalar parameter and c a fixed constant 3-vector, one sees that $\bar{\xi}_i = \xi_i + \epsilon c$ is a family of transformations to which (8), §96 is applicable. But the partial derivative $(\bar{\xi}_i)_\epsilon \equiv c$. Hence, the scalar $\sum c \cdot L_{\xi'_i}$ is, for every constant $c = (c_1, c_2, c_3)$, independent of t along any solution (8) of (5). Choosing successively $(c_1, c_2, c_3) = (1, 0, 0); (0, 1, 0); (0, 0, 1)$, one finds that there exists for every solution (8) of (5) a constant 3-vector A such that $\sum L_{\xi'_i} = A$.

Consequently, (5) has the six scalar integrals

$$(10_1) \quad \sum m_i \xi'_i = A; \quad (10_2) \quad \sum m_i \xi_i - t \sum m_i \xi'_i = B,$$

where the pair of 3-vectors A, B represents six integration constants.

In fact, $\sum L_{\xi_i} = A$ is, by (6₁), equivalent to (10₁); while (10₁) implies that, for some constant 3-vector B ,

$$(11) \quad \sum m_i \xi_i = At + B.$$

Notice that each of the scalar components of the 3-vector equation (11) contains two integration constants, hence cannot represent an integral (§82); but that (11) can be written with the help of the three integrals (10₁) in the form (10₂) of three integrals.

§317 bis. Division of (11) by the total mass $\mu = \sum m_i$ shows that the path of the centre of mass of the n particles in the coordinate system ξ has the equation $\xi = A^*t + B^*$, where the vectors $A^* = \mu^{-1}A$ and $B^* = \mu^{-1}B$ are integration constants. Accordingly, the content of the six integrals (10₁)–(10₂) is that, for any given solution (8) of (5), the motion of the centre of mass is uniform† in the given inertial coordinate system ξ .

§318. The most general Euclidean coordinate transformation (motion) is of the form

$$(12) \quad \bar{\xi} = \Omega \xi + \omega,$$

where the orthogonal matrix Ω of determinant $+1$ and the 3-vector ω represent the rotational and translational component of the motion, respectively, and are arbitrarily given functions of t . It will be assumed that $\Omega = \Omega(t)$, $\omega = \omega(t)$ have continuous second derivatives Ω'' , ω'' .

According to §313, a coordinate system ξ is called inertial if (5) is valid in it. Correspondingly, a transformation (12) will be called inertial if $\bar{\xi}$ is an inertial coordinate system whenever ξ is; i.e., if the pair $\Omega(t)$, $\omega(t)$ is such that (5) and (12) imply the equations which one obtains by writing $\bar{\xi}$ for ξ in (5).

It is clear from (3₂)–(3₃) that, whether the motion (12) is or is not inertial, $U_{\bar{\xi}_i}(\bar{\xi}_1, \dots, \bar{\xi}_n) \equiv \Omega U_{\xi_i}(\xi_1, \dots, \xi_n)$ in virtue of (12); so that $U_{\bar{\xi}_i}(\bar{\xi}_1, \dots, \bar{\xi}_n) = m_i \Omega \xi_i''$, by (5). Consequently, the condition for an inertial transformation is that $\bar{\xi}_i'' = \Omega \xi_i''$ be an identity in t in virtue of (12). If one puts $\Xi = \bar{\xi}_i - \omega$ and $X = \xi_i$, this requirement can be expressed by saying that $\Omega^{-1}\Xi'' + \Omega^{-1}\omega'' = X''$ is an identity in virtue of $\Xi = \Omega X$. But $\Xi = \Omega X$ is the same thing as (8), §69; so that $\Omega^{-1}\Xi''$ is given by (10₂), §69. Consequently, the

† By a uniform motion of a point is meant a rectilinear motion with constant velocity (which may vanish).

transformation (12) determined by a pair $\Omega(t)$, $\omega(t)$ is inertial if and only if

$$(13) \quad 2\Sigma X' + (\Sigma' + \Sigma^2)X + \Omega^{-1}\omega'' = 0$$

is an identity in itself for every t , where $\Sigma = \Sigma(t)$ is the matrix defined by (5), §66. Moreover, $X = X(t)$ was defined as $\xi_i = \xi_i(t)$, where i has one of the values $1, \dots, n$; while the values $\xi_i(t)$, $\xi'_i(t)$ of a solution (8) of (5) can be chosen at any fixed t as arbitrary initial values. Since $\Sigma(t)$; $\Omega(t)$, $\omega(t)$ depend only on the transformation (12) and not on the particular solution (8), it follows that (12) is an inertial transformation if and only if the coefficients 2Σ , $\Sigma' + \Sigma^2$, $\Omega^{-1}\omega''$ of (13) vanish for every t . Clearly, this will be the case if and only if $\Sigma(t) \equiv 0$ and $\omega''(t) \equiv 0$. But $\Sigma(t) \equiv 0$ means, by (5), §66 and the end of §69, that the rotational component of (12) is independent of t ; while $\omega''(t) \equiv 0$ means that the translational component of (12) is a uniform motion (cf. the footnote to §317 bis).

Accordingly, (12) is an inertial transformation if and only if it is of the form

$$(14) \quad \bar{\xi} = \Omega\xi + \alpha t + \beta, \text{ where } \Omega; \alpha, \beta \text{ are independent of } t.$$

Notice that the constant rotation matrix Ω , as well as either of the constant vectors α , β , contains three scalar parameters.

§318 bis. The content of the criterion (14) is that if $\bar{\xi} = \Omega(t)\xi'$ where $\Omega(t) \neq \text{const.}$, the rotation of the coordinate system $\bar{\xi}$ about the origin of the inertial coordinate system ξ spoils the validity of Newton's law (5), since there appear, besides the given Newtonian forces U_{ξ_i} , "apparent" forces which act on m_i and are, in view of (13), represented by the "Coriolis force" $2m_i\Sigma(t)\xi'_i(t)$ and the "centrifugal force" $m_i P(t)\xi_i(t)$, where $P = \Sigma' + \Sigma^2$; and that if $\bar{\xi} = \xi + \omega(t)$, where $\omega'(t) \neq \text{const.}$, the non-uniformity of the rectilinear motion of the origin of the coordinate system $\bar{\xi}$ introduces a similar "apparent" force of "accelerated translation," this force being $m_i\omega''(t)$, by (13). In both cases, the "inertia" of the particles is modified precisely by the use of a coordinate system $\bar{\xi}$ which is not inertial in the sense of §313.

§319. The ten integration constants (7_2) , (9) , (10_1) , (10_2) were defined with reference to a given inertial coordinate system ξ . While these integrals clearly exist also with reference to another inertial coordinate system, $\bar{\xi}$, the transformation (14) can change the con-

stants h , C , A , B into other constants, say \bar{h} , \bar{C} , \bar{A} , \bar{B} . It will be sufficient to study the effect of this change in case of the three generating subgroups

$$(i) \quad \bar{\xi} = \Omega \xi; \quad (ii) \quad \bar{\xi} = \xi + \beta; \quad (iii) \quad \bar{\xi} = \xi + \alpha t$$

of the full group (14) of inertial transformations.

(i) If $\bar{\xi} = \Omega \xi$, where $\Omega = \text{const.}$, then, as verified at the beginning of §316, both (3₁) and (3₂) remain invariant; hence, not only does (2₃) remain invariant but one has also $\bar{h} = h$, by (7₂). On the other hand, $\bar{\xi} \times \bar{\xi}' = (\Omega \xi) \times (\Omega \xi') \equiv \Omega(\xi \times \xi')$, by the definition of $u \times v$; so that $\bar{C} = \Omega C$, by (9). Accordingly, not only does $|\bar{C}| = |C|$ hold but, in addition, the angular momentum vector has (if $C \neq 0$) in the three-dimensional Euclidean space a direction which is independent of the choice of the Cartesian system with a given origin; i.e., C actually is a (Cartesian) vector.

(ii) If $\bar{\xi} = \xi + \beta$, where $\beta = \text{const.}$, then, as verified at the beginning of §317, both (3₁) and (3₂) remain invariant; hence, not only does (2₃) remain invariant but one has also $\bar{h} = h$, by (7₃). On the other hand, $\bar{\xi} \times \bar{\xi}' \equiv (\xi + \beta) \times (\xi + \beta)' \equiv (\xi \times \xi') + (\beta \times \xi')$. Hence, $\bar{C} = C + (\beta \times A)$, by (9) and (10₁). Finally, (10₁) and (10₂) show that $\bar{A} = A$ but $\bar{B} = B + \mu\beta$, where $\mu = \sum m_i$.

(iii) If $\bar{\xi} = \xi + \alpha t$, where $\alpha = \text{const.}$, then, while (3₂) shows that (3₂) remains invariant, (3₁) is not invariant, since it goes over into $\frac{1}{2} \sum m_i (\xi'_i + \alpha)^2$. Thus, the Lagrangian function (2₃) is not invariant, although the Lagrangian equations remain invariant ($\bar{\xi} = \xi + \alpha t$ being, by (14), an inertial transformation). However, the change in (2₃) becomes in virtue of (5) an additive constant; in fact, this change is represented by the difference of $\frac{1}{2} \sum m_i (\xi'_i + \alpha)^2$ and $\frac{1}{2} \sum m_i \xi'^2_i$; a difference which reduces, by (10), to $\alpha \cdot A + \frac{1}{2} \alpha^2 \mu$, where $\mu = \sum m_i$. Correspondingly, $\bar{h} = h + \alpha \cdot A + \frac{1}{2} \alpha^2 \mu$, since U in (7₂) is invariant. Furthermore, since $\bar{\xi} \times \bar{\xi}' \equiv (\xi + \alpha t) \times (\xi' + \alpha)$ is the sum of the three vector products $\xi \times \xi'$, $(\alpha \times \xi')t$, $\xi \times \alpha$, it is seen from (9) that $\bar{C} = C + (B \times \alpha)$, by (10₁) and (11). Finally, (10₁) and (10₂) show that $\bar{B} = B$ but $\bar{A} = A + \mu\alpha$, where $\mu = \sum m_i$.

Notice the parallelism of the transformation formulae of C , B , A in the cases (ii), (iii).

§320. According to (6₁)–(6₃), the Hamiltonian forms of the Lagrangian equations (5) and of their energy integral (7₂) are

$$(15_1) \quad \eta'_i = -H_{\xi_i}, \quad \xi'_i = H_{\eta_i}, \quad \text{where } H = \frac{1}{2} \sum m_i^{-1} \eta_i^2 - U; \quad (15_2) \quad H = h;$$

while, by (6₁), the nine integrals (9), (10₁), (10₂) of (5) can be written as

$$(16_1) \sum \xi_i \times \eta_i = C; \quad (16_2) \sum \eta_i = A; \quad (16_3) \sum m_i \xi_i - t \sum \eta_i = B.$$

The procedure of §92, when applied to the nine integrals (16₁)–(16₃) of (15₁), fails to supply new integrals. This is seen from (30), §24 by observing that the three scalar components of any of the three 3-vectors (16_{*j*}), where *j* = 1, 2, 3, are, save for the notation, identical with the three scalar functions F_{3j-2} , F_{3j-1} , F_{3j} which are defined by (29₁)–(29₂), §24 and occur in (30), §24.

In contradistinction to the angular momentum (2₁), the sum (16₂) of all momenta (6₁) is called the linear momentum. Thus, the seven integrals (15₂), (16₁), (16₂) respectively express the conservation of the energy, angular momentum and linear momentum along any solution; while the three non-conservative integrals (16₃) are, by the end of §317 bis, only another formulation for the conservation of linear momentum or for the uniformity of the motion of the centre of mass.

It is clear from §317–§318 that the nine integrals (16₁)–(16₃) correspond to the nine parameters which occur in the group (14) of all inertial transformations (Ω , α , β each containing three scalar parameters). Similarly, §96 bis shows that the tenth known integral, (15₂), corresponds to the fact that (5), being a conservative system, remains invariant if one replaces *t* by $\bar{t} = t + \text{const.}$ In fact, if *t* is an absolute time in the sense of §313, then \bar{t} is also an absolute time if $\bar{t} = t - t^0$ (and, in view of §160, only if $\bar{t} = \pm t - t^0$), t^0 being an arbitrary constant. The transformation group with ten parameters, which corresponds to the existence of the ten integrals (15₂)–(16₃) and is obtained by adjoining $\bar{t} = t - t^0$ to (14), is usually referred to as the Galilei group.

Since ξ_i and η_i , where *i* = 1, . . . , *n*, are 3-vectors, (15₁) has, by §82, exactly $2 \cdot 3 \cdot n$ independent integrals; while (15₂)–(16₃) represent only $1 + 3 + 3 + 3$ of them for every $n (\geq 2)$. If $n > 2$, none of the missing $6n - 10$ integrals is known (if $n = 2$, the missing 2 integrals can be exhibited; cf. §218 bis). Similarly, (15₁) has, again by §82, exactly $6n - 1$ conservative integrals, while only 7 of them are known, namely (15₂)–(16₂). The situation is understandable from §130 and §199.

§320 bis. In this direction, there should be mentioned a result of Bruns, which states that if $n > 2$, then (15₂)–(16₂) exhaust all those

independent conservative integrals of (15₁) which are algebraic functions of the canonical variables ξ_1, \dots, η_n ; and that, roughly speaking, the same holds if one adjoins to the field of algebraic functions the sign \int , thus allowing Abelian functions of ξ_1, \dots, η_n . Cf., however, §129.

On the other hand, Poincaré has established, for $n \geq 3$, a result which concerns the non-existence of additional isolating (= "uniforme") integrals and is, therefore, not open to the objection of §129. Nevertheless, his result, as well as its formal refinement obtained by Painlevé, is not satisfactory from the point of view of §129–§130. In fact, these negative results do not deal with the case of fixed, but rather with unspecified, values of the masses m_i in (5), and assume, in addition, that the integrals whose existence has to be disproved depend on the variable values of the parameters m_i in a certain analytic manner. Clearly, these assumptions in themselves do not allow any dynamical interpretation, since a dynamical system (5) is determined by a *fixed* set of n positive numbers m_i .

§321. By the problem of n bodies is meant the problem assigned by the system (5) of differential equations, if U is given by (3₂)–(3₃).

The explicit representation (3₂)–(3₃) of the force function was not used above; in fact, §316–§319 can be repeated without any change for every $U(\xi_1, \dots, \xi_n)$ which is invariant under the six-parametric transformation group $\bar{\xi} = \Omega\xi + \omega$ of Euclidean movements. For instance, one can choose U so that the attraction becomes proportional to any fixed, instead of the -2 nd, power of the distance, i.e., one can replace $1/\rho_{jk}$ in (3₂) by $1/\rho_{jk}^\lambda$ or, rather, by $\pm 1/\rho_{jk}^\lambda$, where $\pm \lambda \leq 0$, in order that the forces become attractive, and not repulsive.

There is a simplification in the three cases $\lambda = 0, \pm 2$. For if $\lambda = -2$, then U becomes a quadratic form, and (5) can be separated (by means of linear conservative inertial transformations which are not inertial in Newton's case $\lambda = -1$) into $3n$ linear systems $q'' + aq = 0$ with a single degree of freedom, each of which has an energy integral $\frac{1}{2}(q'^2 + aq^2) = \text{Const.}$ (the constants a are determined by the m_i). If $\lambda = 0$, then every $a = 0$, since (5) becomes $\xi_i'' = 0$. Finally, if $\lambda = 2$, one is dealing with the exceptional case of §161, where β is the present $-\lambda$. In this case of an attraction inversely proportional to the cube of the distance, (5) has, in addition to the ten integrals (15₂)–(16₃), the pair of elementary integrals

$$(17) \quad \begin{aligned} \sum m_i \xi_i' \cdot (\xi_i - t \xi_i') + 2tU &= \text{const.}, \\ \frac{1}{2} \sum m_i (\xi_i - t \xi_i')^2 - t^2 U &= \text{Const.}, \end{aligned}$$

which are indeed the integrals mentioned at the end of §161.

§321 bis. Many† of the results of the following sections will hold not only in Newton's case of gravitation but for almost any value of the exponent λ , and often also for cases in which U is not homogeneous. Unfortunately, this situation must not be looked upon as an expression of the enchanting generality of the results to be obtained, but rather as a drastic manifestation of the fact that practically everything that is known on the problem (5) of n bodies is superficial enough to hold without the explicit assumption (3₂)–(3₃) also.

In this sense, the sharp statement of §217–§218 bis must be considered as a deep result. And the fact is that, if $n > 2$, not a single result, analytical or topological, of comparable sharpness has ever been formulated.

Consequences of the Conservation Integrals

§322. If ξ is any inertial coordinate system, then, by §314–§315,

$$\begin{aligned} (1_1) \quad m_i \xi_i'' &= U_{\xi_i}; & (1_2) \quad \rho_{jk} &= |\xi_j - \xi_k|; \\ (1_3) \quad U &= \sum^* m_j m_k / \rho_{jk}; & (1_4) \quad T &= \frac{1}{2} \sum m_i \xi_i'^2, \\ (2_1) \quad T - U &= h; & (2_2) \quad \mu &= \sum m_i; \\ (2_3) \quad J &= \sum m_i \xi_i^2; & (2_4) \quad J'' &= 2U + 4h. \end{aligned}$$

According to the end of §317 bis, the motion of the centre of mass, $\xi^* = \mu^{-1} \sum m_i \xi_i$, belonging to any given solution

$$(3) \quad \xi_i = \xi_i(t), \quad (i = 1, \dots, n)$$

of (1₁) is a uniform motion with reference to the given inertial coordinate system ξ . Hence, the coordinate transformation $\bar{\xi} = \xi - \xi^*$ is of the form (14), §318, i.e., an inertial transformation. In other words, if ξ is an inertial coordinate system, then so is the barycentric coordinate system $\bar{\xi}$ whose axes are parallel to those of ξ , where it is understood that a barycentric coordinate system with reference to a given solution (3) of (1₁) is defined as one which has the centre of

† Exceptional are, of course, results which are explicit or depend on explicit estimates, e.g., on the inequalities of §343–§344.

mass as origin for every t . Accordingly, (1₁)–(1₄) remain valid if one writes $\bar{\xi}_i = \xi_i - \xi^*$ for ξ_i .

In all that follows, it will always be assumed that the inertial coordinate system ξ has already been chosen as barycentric, i.e., that $\bar{\xi}_i = \xi_i$, where $\bar{\xi}_i = \xi_i - \xi^*$; so that $\mu^{-1} \sum m_i \xi_i$ vanishes for every t . In other words, the inertial coordinate system can, and will, be so chosen that in (11), §317 the six integration constants represented by A, B vanish. Thus, the nine integrals (9)–(10₂), §316–§317 reduce to

$$(4_1) \quad \sum m_i \xi_i \times \xi'_i = C; \quad (4_2) \quad \sum m_i \xi_i = 0; \quad (4_3) \quad \sum m_i \xi'_i = 0,$$

where C denotes the constant angular momentum with reference to the centre of mass, which now is the origin O of the inertial coordinate system ξ . Similarly, (2₁) and (2₃) are the (constant) energy and the (not, in general, constant) polar inertia momentum of (3) with reference to O .

Notice that (4₂)–(4₃) are not integrals but merely invariant relations of (1₁), since the integration constants A, B are particularized to 0 (cf. §82). Furthermore, (4₃) is implied by (4₂), since (4₂) holds for every t .

The projections of the vectors ξ_i, C on the axes $\xi^I, \xi^{II}, \xi^{III}$ of the coordinate system ξ will be denoted by ξ_i^ν, C^ν , respectively, where $\nu = I, II, III$; so that (4₁) represents the three scalar integrals

$$(5) \quad \sum m_i (\xi_i^\alpha \xi_i'^\beta - \xi_i^\beta \xi_i'^\alpha) = C^\gamma,$$

where $(\alpha, \beta, \gamma) = (I, II, III), (II, III, I), (III, I, II)$.

§322 bis. In virtue of (4₂)–(4₃), the sums (2₃) and (1₄) become expressible in terms of the $\frac{1}{2}n(n-1)$ mutual distances $\rho_{jk} = |\xi_j - \xi_k|$ and the $\frac{1}{2}n(n-1)$ mutual speeds $|\xi'_j - \xi'_k|$, respectively; in fact,

$$J = \mu^{-1} \sum^* m_j m_k \rho_{jk}^2; \quad T = \frac{1}{2} \mu^{-1} \sum^* m_j m_k (\xi'_j - \xi'_k)^2,$$

where μ is defined by (2₂), and \sum^* has the same meaning as in (1₃). In order to obtain these representations of (1₄), (2₃) from (4₂), (4₃), it is sufficient to apply the identity (1), §65 to $a_i = m_i^{\frac{1}{2}}$ and to each of the scalar components b_i of the 3-vectors $m_i^{\frac{1}{2}} \chi_i$, where either $\chi_i = \xi_i$ or $\chi_i = \xi'_i$.

It is similarly verified from (4₂)–(4₃) and (2₂) that (4₁) can be written in the form

$$C = \mu^{-1} \sum^* m_j m_k (\xi_j - \xi_k) \times (\xi'_j - \xi'_k).$$

It may be mentioned that if $n = 3$, then, since $\sum m_i \xi_i = 0$,

$$\mu \xi_i = m_k \zeta_j - m_j \zeta_k, \quad \text{where} \quad \zeta_i = \xi_k - \xi_j;$$

$$(i, j, k) = (1, 2, 3), (2, 3, 1), (3, 1, 2).$$

§323. Since the coordinate systems ξ and $\bar{\xi} = \Omega \xi$ have, for every rotation matrix $\Omega = \Omega(t)$, the same origin, $\bar{\xi} = \Omega(t) \xi$ is a barycentric coordinate system whenever ξ is. However, the criterion (14), §318 shows that the barycentric coordinate system $\bar{\xi} = \Omega \xi$ is, for a given inertial barycentric coordinate system ξ , again inertial only when Ω is independent of t ; so that the variety of all inertial barycentric coordinate systems depends only on the three constants which determine an arbitrary rotation matrix $\Omega = \text{const.}$

It will now be shown that these three constants enable one to choose the inertial barycentric coordinate system so that for the integration constants (5), which represent the components of the constant vector (4₁), one has

$$(6) \quad \begin{aligned} C^I &= 0, \quad C^{II} = 0, \quad \text{i.e.,} \\ C^{III} &= \pm \left| (C^I)^2 + (C^{II})^2 + (C^{III})^2 \right|^{\frac{1}{2}} \equiv \pm |C|. \end{aligned}$$

Furthermore, if $C \neq 0$, one can choose the barycentric inertial coordinate system in such a way that

$$(7) \quad C^I = 0, \quad C^{II} = 0, \quad C^{III} = |C|;$$

while if $C = 0$, then (7) holds in every inertial barycentric coordinate system.

First, if ξ and $\bar{\xi} = \Omega \xi$ is any pair of inertial barycentric coordinate systems, and if C, \bar{C} denote the corresponding constant vectors of the angular momentum, then $\bar{C} = \Omega C$, by (i), §319. Hence, $|\bar{C}| = |C|$ and $\bar{C} \cdot \bar{\xi} = C \cdot \xi$. If one excludes, for a moment, the case $C = 0$, the equation $C \cdot \xi = 0$ determines a plane through the origin; and $\bar{C} \cdot \bar{\xi} = C \cdot \xi$ shows that $\bar{C} \cdot \bar{\xi} = 0$ is the equation of the same plane. In other words, the plane which goes through the centre of mass and is perpendicular to the vector of angular momentum is not only independent of t (since $C = \text{const.}$) but is, in addition, one and the same plane in every inertial barycentric coordinate system. This plane, which is defined only if the vector integration constant $C \neq 0$, is called the invariable plane of the given solution (3) of (1₁). It is clear from (4₁) or (5) that (6) holds if and only if the invariable plane is chosen to be the (ξ^I, ξ^{II}) -plane of the barycentric inertial

coordinate system $\xi = (\xi^I, \xi^{II}, \xi^{III})$; and that (7) is satisfied by any barycentric inertial coordinate system whose oriented ξ^{III} -axes is parallel to the normal of the oriented invariable plane, if the positive normal of this plane is defined as having the direction of the angular momentum vector. Finally, if this vector vanishes, i.e., if the invariable plane does not exist, then (6) and (7) hold in every inertial coordinate system, since $\bar{C} = C$ when $C = 0$.

Notice that the energy constant of (3) is the same in every inertial barycentric coordinate system; cf. (i), §319.

§324. A given solution (3) of (1₁) will be called planar if there exists a plane Π^* which contains all n bodies for every t and has, with reference to the barycentric inertial coordinate system ξ , a position which does not depend on t .

It will be shown that if the solution is planar, then Π^* is the invariable plane, provided that the solution has an invariable plane (i.e., if $C \neq 0$).

In fact, it is clear that if Π^* exists, it contains the centre of mass; so that, since Π^* is independent of t , one can choose a barycentric inertial coordinate system ξ such that Π^* becomes the (ξ^I, ξ^{II}) -plane. Then $\xi_i^{III} = \xi_i^{III}(t)$ vanishes for every t and i . Hence, (5) shows that (6) is satisfied. Since (6) is, by §323, the necessary and sufficient condition for a coordinate system ξ whose (ξ^I, ξ^{II}) -plane is the invariable plane in the case $C \neq 0$, the proof is complete.

It is clear from the uniqueness of the initial value problem of the differential equations (1₁), that a solution (3) is planar if and only if there exists for a fixed $t = t_0$ a plane such that not only the n initial position vectors $\xi_i(t_0)$ but also the n initial velocity vectors $\xi'_i(t_0)$ are contained in this plane.

§325. A given solution (3) of (1₁) will be called flat if there exists for every t a plane $\Pi = \Pi(t)$ which contains all n bodies at this t .

Not every flat solution is planar. In fact, while every solution of the problem of $n = 3$ bodies is flat, the last remark of §324 shows that a solution of the problem of $n = 3$ bodies is not, in general, planar. Actually, there exist flat non-planar solutions for every $n > 3$ also.†

† In fact, let $n = 4$ masses m_i be pairwise equal ($m_1 = m_2, m_3 = m_4$), and choose the initial positions and initial velocities of both pairs of equal masses symmetric with respect to one and the same plane P through the centre of mass, O . The position of either pair of equal masses will then be symmetric with respect to P for every t ; so that the $n = 4$ masses are contained for every

§325 bis. With reference to a given flat solution (3), where n is arbitrary, introduce instead of the inertial barycentric coordinate system $\xi = (\xi^I, \xi^{II}, \xi^{III})$ a barycentric (but not, in general, inertial) coordinate system $\bar{\xi} = (x, y, z)$ which rotates about the centre of mass in such a way as to have the plane $\Pi(t)$ of the n bodies as (x, y) -plane for every t ; so that

$$(8_1) \quad \xi = \Omega \bar{\xi}; \quad (8_2) \quad \bar{\xi}_i = (x_i, y_i, z_i); \quad (8_3) \quad z_i(t) \equiv 0, \quad (i = 1, \dots, n),$$

where $\Omega = \Omega(t)$ is a rotation matrix.* Put

$$(9_1) \quad J^{xx} = \sum m_i x_i^2, \quad J^{yy} = \sum m_i y_i^2; \quad J^{xy} = \sum m_i x_i y_i;$$

$$(9_2) \quad K = \sum m_i (x_i y_i' - y_i x_i')$$

and define $S = (s_1, s_2, s_3)$ in terms of Ω by means of (5), §66. It will be shown that

$$(10_1) \quad \Omega^{-1} \begin{pmatrix} 0 \\ 0 \\ |C| \end{pmatrix} = \begin{pmatrix} s_1 J^{yy} - s_2 J^{xy} \\ s_2 J^{xx} - s_1 J^{xy} \\ K + s_3 (J^{xx} + J^{yy}) \end{pmatrix};$$

$$(10_2) \quad \begin{vmatrix} J^{xx} & J^{xy} \\ J^{xy} & J^{yy} \end{vmatrix} = \sum^* m_j m_k \begin{vmatrix} x_j & x_k \\ y_j & y_k \end{vmatrix}^2; \quad (10_3) \quad J^{xx} + J^{yy} = J.$$

First, (8₂), (8₃) show that the components of the vector $\bar{\xi}_i \times \bar{\xi}_i'$ are

t in a plane $\Pi(t)$ which is perpendicular to \mathbf{P} and rotates, in accordance with (4₁), about the normal to \mathbf{P} through O . Hence, the solution is necessarily flat, although it is planar only if the initial angular velocity of the plane $\Pi(t)$ is chosen to be zero.

By increasing the symmetry, one can obtain a flat non-planar solution for any $n > 4$ also. In fact, let the initial positions and initial velocities of four equal masses be selected in such a way as to be symmetric not only with respect to a plane \mathbf{P} but also with respect to a line \mathbf{K} perpendicular to \mathbf{P} , both \mathbf{P} and \mathbf{K} containing O ; moreover, let the initial positions and initial velocities of an arbitrary number ($= [\frac{1}{2}n] - 2$) of pairs of equal masses be chosen in the line \mathbf{K} so as to be symmetric with respect to \mathbf{P} ; finally, if n is odd, let in addition an arbitrary mass with initial velocity zero be placed at the intersection of \mathbf{P} and \mathbf{K} , i.e., at O . Then neither the planar symmetry nor the axial symmetry can be disturbed by the resulting motion. Thus, again all masses will be contained for every t in a plane $\Pi(t)$ through the line \mathbf{K} ; so that the solution is flat, although not, in general, planar.

* Notice that $\Omega(t)$ is not uniquely determined by (8₃), since the position of the x -axes can be chosen arbitrarily within the plane $\Pi(t)$. For instance, one can normalize the rotation by requiring that m_1 be on the x -axes for every t , in which case $\Omega(t)$ becomes an analytic function of t (the reason being that every solution (3) of the analytic differential equations (1₁) is analytic).

0, 0, $(x_i y_i' - y_i x_i')$; while those of $\bar{\xi}_i \times (S \times \bar{\xi}_i)$ are $s_1 y_i^2 - s_2 x_i y_i$, $s_2 x_i^2 - s_1 x_i y_i$, $s_3(x_i^2 + y_i^2)$, those of S being s_1, s_2, s_3 . Hence, it is seen from (9₁)–(9₂) that the vector on the right of (10₁) is the sum of $\sum m_i \bar{\xi}_i \times \bar{\xi}_i'$ and $\sum m_i \bar{\xi}_i \times (S \times \bar{\xi}_i)$. On the other hand, it is assumed that the barycentric inertial coordinate system $\xi = (\xi^I, \xi^{II}, \xi^{III})$ is chosen in accordance with (7); so that the vector on the left of (10₁) is $\Omega^{-1}C \equiv \sum m_i \Omega^{-1}(\xi_i \times \xi_i')$, by (5), (4₁). Consequently, the statement (10₁) is true if $\Omega^{-1}(\xi_i \times \xi_i')$ is the sum of $\bar{\xi}_i \times \bar{\xi}_i'$ and $\bar{\xi}_i \times (S \times \bar{\xi}_i)$. This, when combined with (11₂), §69, completes the proof of (10₁), §325, since (8₁), §325 becomes identical with (8), §69, if one puts $\xi = \Xi, \bar{\xi} = X$.

Next, (4₁)–(4₂), §314 and the definitions (9₁), §325 show that (10₂), §325, is merely the particular case $a_i = m_i^{\frac{1}{2}}x_i, b_i = m_i^{\frac{1}{2}}y_i$ of (1), §65.

Finally, $\xi_i^2 = \bar{\xi}_i^2 = x_i^2 + y_i^2 + z_i^2 = x_i^2 + y_i^2$, by (8₁)–(8₃); so that (10₃) is clear from (9₁) and (2₃).

§326. Trivial examples show* that there exist solutions which are neither flat nor such as to possess an invariable plane. On the other hand, a solution for which the invariable plane does not exist (i.e., for which $C = 0$) is necessarily planar (§324) whenever it is flat (§325).

In order to prove this, notice that if $C = 0$, then the relation (10₁), which is valid for every flat solution, implies for $s_1 = s_1(t), s_2 = s_2(t)$ two homogeneous linear equations which have (10₂) as determinant. Since, by the footnote to §325 bis, this determinant, as well as s_1, s_2 , can be considered as analytic in t , it follows that either s_1 and s_2 vanish for every t or (10₂) does.

In the first case, where $s_1(t) \equiv 0 \equiv s_2(t)$, the condition (13₂), §72 is satisfied, and so the z -axis of the rotating coordinate system $\bar{\xi} = (x, y, z)$ coincides with the ξ^{III} -axis of the inertial coordinate system $\xi = (\xi^I, \xi^{II}, \xi^{III})$. This, when combined with (8₃), shows that all n bodies move within the fixed (ξ^I, ξ^{II}) -plane; so that the solution is planar.

In the second case, where the determinant (10₂) vanishes for every t , so does each of the $\frac{1}{2}n(n-1)$ determinants $x_i y_k - x_k y_i$, where $1 \leq i < k \leq n$. It follows, therefore, from (8₂) that the area of the triangle formed by the origin and any two of the n masses vanishes

* For instance, let n be one of the numbers 4, 6, 8, 12, 20, and place n equal masses m_i at the vertices of a regular solid, choosing the initial velocities so as to have a common magnitude and to be directed towards the mid-point of the regular solid.

identically; and so the n masses are collinear for every t . Hence, the rotating coordinate system (x, y, z) of §325 bis can be chosen so that all n masses are on the x -axis for every t . Then not only (8₃) holds but also $y_i(t) \equiv 0$. Consequently, $J^{\nu\nu} \equiv 0$, $J^{x\nu} \equiv 0$; $K \equiv 0$ and $J^{xx} \equiv J$, by (9₁); (9₂) and (10₃). Hence, (10₁), where $C = 0$ by assumption, reduces to $0 = 0$, $0 = s_2 J$, $0 = s_3 J$. Since (2₃) is positive, it follows that $s_2(t) \equiv 0 \equiv s_3(t)$. This is, save for the notation, identical with the condition $s_1(t) \equiv 0 \equiv s_2(t)$ of the first case, treated before by means of (13₂), §72; so that the proof is complete.

§327. The n masses m_i are said to be in syzygy at a given date $t = t_0$ if they are collinear at this date, where it is understood that the solution (3) under consideration need not be such that the n masses are collinear (or, for that matter, co-planar) when $t \neq t_0$.

It will be shown that at the date of a syzygy all n masses must lie in the invariable plane, provided that there exists an invariable plane (i.e., if $C \neq 0$).

The assumption is that, if ξ_i denotes the position of m_i at $t = t_0$ in the barycentric inertial coordinate system ξ , then any pair of the n points ξ_i is collinear with the origin; so that $\xi_i \times \xi_k = 0$, where $i, k = 1, \dots, n$. Hence, $(\xi_i \times \xi_k) \cdot \xi'_i = 0$, i.e., $(\xi_i \times \xi'_i) \cdot \xi_k = 0$; and so scalar multiplication of (4₁) by ξ_k gives $0 = C \cdot \xi_k$, where $k = 1, \dots, n$. Since $C \cdot \xi = 0$ is, if $C \neq 0$, the equation of the invariable plane, the proof is complete.

§328. A given solution (3) of (1₁) will be called rectilinear if there exists a line Λ^* which contains all n bodies for every t and has, with reference to the inertial coordinate system $\xi = (\xi^I, \xi^{II}, \xi^{III})$, a position which is independent of t .

It will be shown that in this case the solution (3) cannot exist for $-\infty < t < +\infty$ without leading, at some finite $t = t^0$, to a collision of at least two of the n bodies.

First, let Λ^* be chosen as the ξ^I -axis, so that $\xi_i = (\xi_i^I, 0, 0)$, and let the numeration of the m_i be chosen in such a way that $\xi_1^I < \xi_2^I < \dots < \xi_n^I$. Then, since the centre of mass is the origin, ξ_n^I is positive; while (1₁), (1₂), (1₃) show that the second derivative of ξ_n^I is negative (in fact, m_1, m_2, \dots, m_{n-1} attract m_n in the direction of $\xi^I = -\infty$). This, when applied for $-\infty < t < +\infty$, contains a contradiction, since it is impossible to draw in a (t, f) -plane (where $f = \xi_n^I$) a curve $f = f(t)$ which runs within the upper half-plane ($f > 0$) and is concave from below ($f'' < 0$) for $-\infty < t < +\infty$. The contradiction can be removed only by assuming either

that the solution (3) does not exist for $-\infty < t < +\infty$ but only on a limited t -range or that there is a collision at some finite $t = t^0$ (in which case at least one of the denominators (1_2) of (1_3) vanishes, and so (1_1) becomes illusory).

§329. A given solution (3) of (1_1) will be called collinear if there exists for every t a line $\Lambda = \Lambda(t)$ which contains all n bodies at this t .

While this obviously implies that the solution is flat (§325), it does not imply that the solution is rectilinear (§328), since the line $\Lambda(t)$ is allowed to vary with t . However, $\Lambda(t)$ must then rotate about the centre of mass in such a way as to be contained in a plane Π^* which has a fixed position with reference to the barycentric inertial coordinate system ξ . In other words, every collinear solution is planar. This is clear from the result of §327 or of §326 according as $C \neq 0$ or $C = 0$.

§330. Incidentally, a collinear solution does not or does have an invariable plane according as the solution is or is not rectilinear. For if the line $\Lambda(t)$ is independent of t , then, since it contains $\xi_i = \xi_i(t)$, it also contains $\xi'_i = \xi'_i(t)$; so that $\xi_i \times \xi'_i = 0$, and so $C = 0$, by (4_1) . If, on the other hand, the line $\Lambda(t)$ is not independent of t , i.e., if the angular velocity of its movement within the plane Π^* does not vanish identically, then $C \neq 0$, as will be seen at the end of §331 bis.

§331. It will be shown that if a collinear solution is not rectilinear, then the geometrical configuration formed by the n masses remains similar to itself when t varies; so that all the mutual distances vary in the same proportion, if they vary at all.

First, the solution is planar, by §329. Hence, the plane Π^* of the movement can be chosen as the (ξ^I, ξ^{II}) -plane of the barycentric inertial coordinate system ξ . Choose in this plane a coordinate system (x, y) which has the same origin as (ξ^I, ξ^{II}) but rotates, with reference to (ξ^I, ξ^{II}) , with a certain angular velocity $\phi' = \phi'(t)$ in such a way that the x -axis is the line $\Lambda(t)$ for every t . Then the coordinate $y_i = y_i(t)$ of every m_i vanishes for every t . Hence, the projection of the absolute acceleration of m_i on the y -axis of the rotating coordinate system, a projection represented by the second line of (14_2) , §73 (where $x = x_i$, $y = y_i$), is seen to be $2\phi'x'_i + \phi''x_i$. On the other hand, all n masses lie on the rotating x -axis; so that the forces of gravitation, i.e., the vectors U_{ξ_i} occurring in (1_1) , have on the y -axis projections which vanish identically. Consequently, $2\phi'x'_i + \phi''x_i$

$= 0$ for every i and t . But the angular velocity $\phi' = \phi'(t)$ is an analytic function of t , and so it vanishes for isolated values of t , at most. In fact, $\phi'(t) \equiv 0$ is excluded by the assumption that the solution is not rectilinear, i.e., that the line $\Lambda(t)$ is actually rotating. Furthermore, $x_i \neq 0$ for at least $n - 1$ of the n values of i . Hence, one can divide $2\phi'x'_i + \phi''x_i = 0$ by $\phi'x_i$ for at least $n - 1$ values of i . It follows, therefore, by a quadrature that $x_i(t) = \lambda(t)x_i(0)$ holds for at least $n - 1$ of the n values of i and for a function $\lambda(t)$ which, being determined by $\phi' = \phi'(t)$ alone, is independent of these i .

In order to complete the proof, one has merely to substitute this into (4₂). In fact, it then becomes clear that the n -th i need not be excluded, i.e., that $x_i(t) = \lambda(t)x_i(0)$ holds for all n values of i .

§331 bis. On inserting this and $y_i(t) \equiv 0$ into the linear equations which define the rotating coordinate system (x, y) in terms of the non-rotating coordinate system (ξ^I, ξ^{II}) and of $\phi = \phi(t)$, and substituting the resulting representation of $\xi_i^I(t)$ and $\xi_i^{II}(t)$ into the definition (5) of C^{III} , one sees from (7) that $|\phi'| \lambda^2 J_0 = |C|$, where J_0 denotes the positive constant $\sum m_i \{x_i(0)^2 + y_i(0)^2\}$. Since obviously $\lambda^2 \neq 0$, it follows that $\lambda(t) = \text{const.}$ if and only if $\phi'(t) = \text{Const.}$ In other words, not only the shape but also the size of the configuration of the n bodies is independent of t if and only if so is the angular velocity $\phi' (\neq 0)$ of the line $\Lambda(t)$.

The relation $|\phi'| \lambda^2 J_0 = |C|$, where $\lambda^2 > 0$, $J_0 > 0$ and $\phi' \neq 0$, implies also that $C \neq 0$, as stated at the end of §330.

§332. The results of §322 and §323–§331 bis are, in the main, consequences of the nine integrals found in §317 and §316, respectively.

As an application of the tenth known integral, i.e., of (2₁) or of its equivalent formulation (2₄), it will be shown that if a given solution (3) of (1₁) exists for all t and is such that the $\frac{1}{2}n(n - 1)$ mutual distances between the n masses remain less than a sufficiently large constant, then the energy constant h must be negative; and that the vis viva, i.e., $2T = 2T(t)$, must then oscillate about the force function $U = U(t)$, in the sense that*

$$(10) \quad \underline{\lim} 2T \leq \underline{\lim} U \leq \overline{\lim} U \leq \overline{\lim} 2T (\leq +\infty), \text{ as } t \rightarrow \infty.$$

Since the centre of mass is $\xi = 0$, the assumption that the mutual

* Either both signs $\underline{\lim}$, $\overline{\lim}$ of lower and upper limits refer to $t \rightarrow -\infty$ or both to $t \rightarrow +\infty$.

distances $\rho_{jk} = |\xi_j - \xi_k|$ remain bounded, as $t \rightarrow \infty$, is clearly equivalent to the assumption that the positive function (2₃) of t remains bounded as $t \rightarrow \infty$; so that $0 \leq \underline{\lim} J \leq \overline{\lim} J < +\infty$. Thus, the ratio $J:t^2$ is neither less than a negative, nor greater than a positive, constant for all sufficiently large t^2 . Hence, application of two quadratures to the second derivative $J'' = J''(t)$ shows that $\underline{\lim} J'' \leq 0 \leq \overline{\lim} J''$. This is, by (2₄), equivalent to

$$\underline{\lim} U \leq -2h \leq \overline{\lim} U,$$

and so, by (2₁), to the statement (10).

Notice that both (1₃) and (1₄) are positive. Since $\underline{\lim} U \leq -2h$, where $U > 0$, it follows also that either $h < 0$ or $\underline{\lim} U = 0$. Hence, the statement $h < 0$ follows from the fact that, by (1₃), the mutual distances become arbitrarily large (for large t) in the case $\underline{\lim} U = 0$.

§332 bis. It is clear from this proof that if $h \geq 0$, then not only $\overline{\lim} J = +\infty$ but also $\lim J = +\infty$.

It is not stated (and it is, if $n > 2$, not true) that the necessary condition $h < 0$ for $\overline{\lim} J < +\infty$ is sufficient as well.

Simultaneous Collisions

§333. Since (2₄) contains both functions J, U of the time, it would be desirable (cf. §332) to have a simple differential equation which contains only one of the functions J, U and its derivatives. Unfortunately, it is not possible to find such an equation. On the other hand, there are several useful inequalities connecting J (or U) and its derivatives.

For instance, there exist two positive constants M_0, m_0 which depend only on the masses and are such that $J = J(t)$ and its derivatives satisfy the inequalities

$$(11_1) \quad |J'''| \leq M_0(|J''| + 4|h|)^{\frac{1}{2}}; \quad (11_2) \quad (J'' - 4h)J^{\frac{1}{2}} \geq m_0 > 0$$

for all solutions (3) of (1₁) which have the arbitrary fixed value $h(\geq 0)$ as energy constant.

The proof of these facts will be based on the relations

$$(12_1) \quad T = \frac{1}{2}\mu^{-1} \sum^* m_j m_k (\xi'_j - \xi'_k)^2; \quad (12_2) \quad J = \mu^{-1} \sum^* m_j m_k \rho_{jk}^2,$$

which were proved in §322 bis.

States of collision are, of course, excluded, i.e., the distances $\rho_{jk} = |\xi_j - \xi_k|$ do not vanish; so that application of the last line of §65 to $v = \xi_j - \xi_k$ shows that ρ'_{jk} exists and $|\rho'_{jk}| \leq |\xi'_j - \xi'_k|$. Hence, from (1₃),

$$|U'| = \left| -\sum^* m_j m_k \rho'_{jk} / \rho_{jk}^2 \right| \leq \sum^* m_j m_k |\xi'_j - \xi'_k| / \rho_{jk}^2,$$

$$\text{where } m_j m_k / \rho_{jk} < U,$$

again by (1₃); so that $|U'| \leq U^2 \sum^* |\xi'_j - \xi'_k| / (m_j m_k)$. Since it is clear from (12₁) that $m_j m_k (\xi'_j - \xi'_k)^2 \leq 2\mu T$, it follows, by placing $M_0 = \mu^{\frac{1}{2}} \sum^* (m_j m_k)^{-\frac{1}{2}}$, that $|U'| \leq M_0 U^2 (2T)^{\frac{1}{2}}$. This completes the proof of (11₁), since (2₄) and (2₁) imply that $U' = \frac{1}{2} J'''$ and $U^2 (2T)^{\frac{1}{2}} \leq \frac{1}{4} (|J''| + 4|h|)^{\frac{1}{2}}$.

Similarly, every term of the sum (12₂) is less than the sum (12₂), i.e., $J^{\frac{1}{2}} m_j m_k / \rho_{jk} > \mu^{-\frac{1}{2}} (m_j m_k)^{\frac{1}{2}}$. Since $J'' - 4h = 2 \sum^* m_j m_k / \rho_{jk}$ by (2₄) and (1₃), it follows that (11₂) is satisfied by $m_0 = 2\mu^{-\frac{1}{2}} \sum^* (m_j m_k)^{\frac{1}{2}}$.

§334. A further limitation of $J = J(t)$ is expressed by

$$(13) \quad J'' - 2h - \frac{1}{4} J'^2 / J \geq C^2 / J,$$

an inequality more elaborate than (11₁)–(11₂), since (13) contains, besides the energy constant h , the length $|C|$ of the constant angular momentum of the arbitrary solution (3).

In order to prove (13), notice first that $\xi_i^2 = |\xi_i|^2$, $|\xi_i \cdot \xi'_i| = |\xi_i| |\xi'_i|$. Hence, the definition $J = \sum m_i \xi_i^2$ implies that $\frac{1}{2} J' = \sum m_i |\xi_i| |\xi'_i|$; and so an application of the inequality $(\sum a_i b_i)^2 \leq (\sum a_i^2)(\sum b_i^2)$ to $a_i = m_i^{\frac{1}{2}} |\xi_i|$, $b_i = m_i^{\frac{1}{2}} |\xi'_i|$ gives

$$\frac{1}{4} J'^2 \leq J \sum m_i |\xi_i|^2 \equiv J \sum m_i (\xi_i \cdot \xi'_i)^2 / \xi_i^2.$$

Similarly, if one puts $a_i = m_i^{\frac{1}{2}} |\xi_i|$ and $A_i = m_i^{\frac{1}{2}} (\xi_i \times \xi'_i) / |\xi_i|$, the definition $C = \sum m_i \xi_i \times \xi'_i$ can be written in the form $C = \sum a_i A_i$; so that

$$C^2 \leq (\sum a_i^2)(\sum |A_i|^2) \equiv J \sum m_i (\xi_i \times \xi'_i)^2 / \xi_i^2.$$

Hence, by addition, $\frac{1}{4} J'^2 + C^2 \leq J \sum m_i \{ (\xi_i \cdot \xi'_i)^2 + (\xi_i \times \xi'_i)^2 \} / \xi_i^2$. But $\{ \} = \xi_i^2 \xi_i'^2$, by (2), §65; so that $\frac{1}{4} J'^2 + C^2 \leq J \sum m_i \xi_i'^2$. This inequality completes the proof of (13), since $\sum m_i \xi_i'^2 = J'' - 2h$, by (1₄), (2₁), (2₄).

§334 bis. If $Q = Q(t)$ denotes the function

$$(14) \quad Q = -2hJ^{\frac{1}{2}} + (\frac{1}{4} J'^2 + C^2) / J^{\frac{1}{2}}, \quad \text{where } J^{\frac{1}{2}} > 0,$$

then, since $(J^{\frac{1}{2}})' = \frac{1}{2} J' / J^{\frac{1}{2}}$, differentiation of (14) shows that Q' is the product of $(J^{\frac{1}{2}})'$ and of a expression which is non-negative in virtue of (13); so that the content of (13) is that Q' and $(J^{\frac{1}{2}})'$ are never of opposite sign. This means that Q and $J^{\frac{1}{2}}$, hence also Q and

J , vary with t in such a way that when either of these functions is increasing, the other cannot decrease.

§335. As an application of this fact, that is, of (13), it will be shown that if a solution (3) of (1₁) has an invariable plane, i.e., if the vector integration constant C does not vanish, then there cannot exist a date $t = t^0$ at which all n masses collide simultaneously.

The simultaneous collision of the n bodies at $t = t^0$ is meant in the sense that all n points $\xi_i = \xi_i(t)$ of the space $\xi = (\xi^I, \xi^{II}, \xi^{III})$ tend, as $t \rightarrow t^0$, to one and the same point of this space; a point which must, of course, be the centre of mass, i.e., the origin $\xi = 0$. Clearly, this will be the case if and only if the positive function $\sum m_i \xi_i^2 = J = J(t)$ tends, as $t \rightarrow t^0$, to 0. Hence, the statement to be proved is that the vanishing of the integration constant $|C|$ of (3) is a necessary condition for the existence of a t^0 such that $J \rightarrow 0$ as $t \rightarrow t^0$.

First, if $J = \sum m_i \xi_i^2$ tends to 0, then so do all the mutual distances $\rho_{jk} = |\xi_j - \xi_k|$; so that the force function $U = \sum^* m_j m_k / \rho_{jk}$ tends to $+\infty$. This means in view of $J'' = 2U + 4h$, where $h = \text{const.}$, that $J'' \rightarrow +\infty$ as $t \rightarrow t^0$. Consequently, J'' is ultimately* positive; hence, J' is ultimately increasing, which implies that, ultimately, J' does not change its sign. Since $0 < J \rightarrow 0$, it follows that, ultimately, J is steadily decreasing. It follows, therefore, from §334 bis that, ultimately, the function (14) is monotone non-increasing. Consequently, the function (14) tends, as $t \rightarrow t^0$, to a limit which might be $-\infty$ but cannot be $+\infty$. On the other hand, this limit of (14) is

$$(15) \quad \lim (\tfrac{1}{4}J'^2 + C^2)/J^{\frac{1}{2}}, \quad (t \rightarrow t^0),$$

since $-2hJ^{\frac{1}{2}} \rightarrow 0$ in view of $h = \text{const.}$ and of $J \rightarrow 0$. But $J^{\frac{1}{2}} > 0$; so that (15) is a finite non-negative limit. This implies that $C^2/J^{\frac{1}{2}}$ must remain bounded as $t \rightarrow t^0$, i.e., as $J \rightarrow 0$; so that, since $C^2 = \text{const.}$, the proof of $C = 0$ is complete.

It was shown at the beginning of this proof that $J'' \rightarrow +\infty$. Hence, (11₁), where M_0, h are constants, implies that, ultimately, $|J'''| < \text{const. } (J'')^{\frac{1}{2}}$.

§335 bis. As an application of (11₂), it will be shown that $J'^2/J^{\frac{1}{2}}$

* That is, when t is sufficiently near to its limit t^0 . Since the problem is reversible (§314), it may be assumed without loss of generality that t tends increasingly to t^0 .

tends to a finite positive limit, when t tends to the date t^0 of a simultaneous collision, i.e., when $J^{\frac{1}{2}} \rightarrow 0$.

As shown in §335, there exists a finite limit (15) which cannot be negative; so that, since $C = 0$, there certainly exists a finite non-negative $\lim J'^2/J^{\frac{1}{2}}$. But the point is that the number $\lim J'^2/J^{\frac{1}{2}}$ cannot be 0.

First, since $C = 0$ and $0 < J^{\frac{1}{2}} \rightarrow 0$, the function (14) and its limit (15) reduce to

$$(16_1) \quad Q = -2hJ^{\frac{1}{2}} + \frac{1}{4}J'^2/J^{\frac{1}{2}}; \quad (16_2) \quad \mu_0 = \frac{1}{4} \lim J'^2/J^{\frac{1}{2}},$$

where $\mu_0 = \lim Q$. It is clear from (16₁) that $(2QJ^{\frac{1}{2}})' = (J'' - 4h)J'$; hence, on integrating $(2QJ^{\frac{1}{2}})'$ between t and some \bar{t} , keeping t fixed and letting \bar{t} tend to the date t^0 of the simultaneous collision, one sees from $\lim J^{\frac{1}{2}} = 0$ and from the finiteness of the limit (16₂) that

$$2QJ^{\frac{1}{2}} = \int_{t^0}^t (J'' - 4h)J'd\bar{t},$$

where $2QJ^{\frac{1}{2}}$ and the integrand belong to the dates t and \bar{t} , respectively. Since, as shown above (§335), the derivative J' ultimately does not change its sign, it follows that the positive constant occurring in (11₂) is such that, ultimately,

$$2|Q|J^{\frac{1}{2}} \geq \int_{t^0}^t (m_0/J^{\frac{1}{2}})J'd\bar{t}.$$

But the last integral is $2m_0J^{\frac{1}{2}}$, since $J^{\frac{1}{2}} \rightarrow 0$ as $\bar{t} \rightarrow t^0$; so that, ultimately, $2|Q|J^{\frac{1}{2}} \geq 2m_0J^{\frac{1}{2}}$, i.e., $|Q| \geq m_0$. Since m_0 is a positive constant and the limit (16₂) of (16₁) exists, the proof of $0 < \lim J'^2/J^{\frac{1}{2}}$ is complete.

§336. Since the number (16₂) does not vanish, it follows that, as $t \rightarrow t^0$, the decreasing function $J = J(t) > 0$ tends to 0 in such a way as to become asymptotically proportional to $(t - t^0)^{\frac{2}{3}}$, with $(\frac{3}{4}\mu_0)^{\frac{1}{3}}$ as factor of proportionality; and that this asymptotic relation remains valid on differentiation with respect to t . In other words,

$$(17_1) \quad J \sim (\tfrac{3}{2}\mu_0^{\frac{1}{3}})^{\frac{2}{3}}(t - t^0)^{\frac{2}{3}}; \quad (17_2) \quad J' \sim (12\mu_0^2)^{\frac{1}{3}}(t - t^0)^{\frac{1}{3}},$$

(where $f_1 \sim f_2$ means that $f_1/f_2 \rightarrow 1$ as $t \rightarrow t^0$, i.e., as $J \rightarrow 0$).

In fact, (17₂) follows from (17₁) not only by formal (that is, unjustified) differentiation but is, in view of (16₂), actually implied by

(17₁). On the other hand, (17₁) follows by writing (16₂) in the form $\pm dt/dJ \sim \frac{1}{2}\mu_0^{-\frac{1}{2}}J^{-\frac{1}{2}}$ and then integrating between $J = 0$ and a nearby $J > 0$; the integration (but not the differentiation) of such an asymptotic formula being clearly always legitimate.

For instance, if $f(\tau)$ tends, as $\tau \rightarrow 0$, to a limit, then so does the average $\tau^{-1} \int_0^\tau f(\sigma) d\sigma$; while the converse is not true, unless f satisfies certain additional conditions. In the theory of limit processes, such additional conditions are called Tauberian conditions.

§337. It will now be shown that, besides (16₂), (17₁), (17₂), one has

$$(18_1) \quad \mu_0 = \lim J^{\frac{1}{2}} J''; \quad (18_2) \quad J'' \sim (\tfrac{2}{3}\mu_0)^{\frac{2}{3}}(t - t^0)^{-\frac{2}{3}}.$$

Notice that (18₁) may be obtained by formal (that is, unjustified) differentiation of (16₂); so that the statement (18₁) implies a refinement of (16₂). Furthermore, it is clear from (17₁) that (18₁) is equivalent to (18₂). Finally, (18₂) manifestly is the result of formal differentiation of (17₂), a process which is legitimate only if a "Tauberian condition" is satisfied. Now, it will be shown the estimate mentioned at the end of §335 supplies such a Tauberian condition.

§338. First, on multiplying (13), where $h = \text{const.}$, by $J^{\frac{1}{2}}$, then letting $t \rightarrow t^0$, $J^{\frac{1}{2}} \rightarrow 0$ and using (16₂), one sees that the lower limit $\underline{\lim} J^{\frac{1}{2}} J'' \geq \mu_0$. Since (18₂) is equivalent to (18₁), it follows that (18₁)-(18₂) will be proved by showing that the upper limit $\overline{\lim} J^{\frac{1}{2}} J'' \leq \mu_0$.

Next, put $F = J'^3$; so that $F'' = 6J'J''^2 + 3J'^2J'''$. On estimating J''' by the inequality $|J'''| < \text{const.} (J'')^{\frac{1}{2}}$, found at the end of §335, and then expressing J' and J'' in terms of $F = J'^3$ and $F' = 3J'^2J''$, one sees that $|F''| < \text{Const.}(F'^2 + |F'|^{\frac{3}{2}})/|F|$. Hence, from (17₂), where $J' = F^{\frac{1}{3}}$,

$$(19) \quad |F''| < \text{const.}(F'^2 + |F'|^{\frac{3}{2}})/|t - t^0|, \quad \text{as } t \rightarrow t^0.$$

Finally, if ν_0 denotes the positive constant $12\mu_0^2$, then

$$(20_1) \quad F \sim \nu_0(t - t^0); \quad (20_2) \quad \underline{\lim} F' \geq \nu_0,$$

as $t \rightarrow t^0$. In fact, (20₁) is the same thing as (17₂), since $F = J'^3$; while comparison of $\nu_0 = 12\mu_0^2$, $F = J'^3$, $F' = 3J'^2J''$ with (16₂) shows that the inequality $\underline{\lim} J^{\frac{1}{2}} J'' \geq \mu_0$, which was proved before, may be written in the form (20₂). Similarly, the inequality $\overline{\lim} J^{\frac{1}{2}} J'' \leq \mu_0$, which remains to be proved, may be written in the form $\overline{\lim} F' \leq \nu_0$. Accordingly, what has to be proved is that

(20₁)–(20₂) and the “Tauberian” condition (19) together imply that $\overline{\lim} F' \leq \nu_0$. (Correspondingly, (20₂) then becomes $F' \rightarrow \nu_0$, which means that the formal differentiation of (20₁) is legitimate).

§338 bis. In order to prove that $\overline{\lim} F' \leq \nu_0$, suppose, if possible, that $\overline{\lim} F' > \nu_0$. The latter inequality, when combined with (20₂), clearly implies the existence of a sequence of t -intervals $t_1^I < t < t_1^{II}, \dots, t_k^I < t < t_k^{II}, \dots$ which tend,* as $k \rightarrow +\infty$, to t^0 in such a way that

$$(21) \quad 0 < \nu_0 < \alpha = F'(t_k^I) < F'(t) < F'(t_k^{II}) = \beta \text{ whenever } t_k^I < t < t_k^{II},$$

where α, β are suitably chosen fixed numbers, situated between the ultimate oscillation limits $\underline{\lim} F', \overline{\lim} F' (\leq +\infty)$ of the continuous function $F'(t)$.

Clearly, one can assume that t^0 is the origin 0. Then on placing $\text{Const.} = \text{const.} (\beta^2 + \beta^{\frac{1}{2}})$, one sees from (19) and (21) that $|F''(t)| < \text{Const.}/|t|$ for any t contained in any of the intervals $t_k^I < t < t_k^{II}$. Since t tends either decreasingly or increasingly to $t^0 = 0$ (cf. the footnote to §335), all t_k^I and t_k^{II} lie on the same side of $t^0 = 0$. Hence, integration of the last estimate of $F''(t)$ between t_k^I and t_k^{II} gives $|F'(t_k^{II}) - F'(t_k^I)| < \text{Const.} \log |t_k^{II}/t_k^I|$. Since $F'(t_k^{II}) - F'(t_k^I)$ is, by (21), the positive constant $\beta - \alpha$, it follows that $\log |t_k^{II}/t_k^I|$ exceeds a positive lower bound, as $k \rightarrow +\infty$. This means that there exists a fixed bound λ such that, as $k \rightarrow +\infty$,

$$(22) \quad \frac{|t_k^{II}|}{|t_k^I|} > \lambda > 1. \text{ Furthermore, } \frac{|F(t_k^I)|}{|t_k^I|} \rightarrow \nu_0, \frac{|F(t_k^{II})|}{|t_k^{II}|} \rightarrow \nu_0,$$

by (20₁), since $t_k^I \rightarrow t^0, t_k^{II} \rightarrow t^0$ and $t^0 = 0, \nu_0 > 0$.

Moreover,

$$(23) \quad \left| \frac{|F(t_k^{II})|}{|t_k^{II}|} - \frac{|F(t_k^I)|}{|t_k^I|} \right| > \alpha \left| \frac{|t_k^{II}|}{|t_k^I|} - 1 \right|,$$

if k is sufficiently large. In fact, all t_k^I, t_k^{II} lie on the same side of t^0 ; so that $||t_k^{II}| - |t_k^I|| = t_k^{II} - t_k^I$, since $t_k^I < t_k^{II}$. Furthermore, since $t_k^I \rightarrow t^0, t_k^{II} \rightarrow t^0$, one sees from (20), where $\nu_0 > 0$, that, if k is sufficiently large, all $F(t_k^I), F(t_k^{II})$ are of the same sign. Hence, the inequality (23), which may clearly be written in the form

* That is, both end points t_k^I, t_k^{II} of these intervals tend to the date t^0 , as $k \rightarrow +\infty$.

$$||F(t_k^{\text{II}})| - |F(t_k^{\text{I}})|| > \alpha ||t_k^{\text{II}}| - |t_k^{\text{I}}||,$$

is equivalent to the inequality $|F(t_k^{\text{II}}) - F(t_k^{\text{I}})| > \alpha(t_k^{\text{II}} - t_k^{\text{I}})$. But the latter inequality is obvious, since, by (21), one has $F'(t) > \alpha > 0$ for $t_k^{\text{I}} < t < t_k^{\text{II}}$. This proves (23).

Letting $k \rightarrow +\infty$ in (23) and using (22), where $\nu_0 > 0$, one obtains $\nu_0|\lambda - \nu_0\nu_0^{-1}| \geq \alpha|\lambda - 1|$. But $|\lambda - \nu_0\nu_0^{-1}| = \lambda - 1 > 0$, by (22), so that $\nu_0 \geq \alpha$; while $\alpha > \nu_0$, by (21). Since the last two inequalities are contradictory, the supposition $\lim F' > \nu_0$, made after the end of §338, is now disproved. Thus, the end of §336 shows that the proof of (18₁)–(18₂) is complete.

§339. It may be mentioned that no solution (3) of (1₁) can approach a state of simultaneous collision when $t \rightarrow \infty$ (where ∞ denotes either $+\infty$ or $-\infty$). In other words, it is impossible that $J \rightarrow 0$ as $t \rightarrow \infty$. Indeed, the proof which was given in §335 for the fact that $J \rightarrow 0$ necessitates $J'' \rightarrow +\infty$, clearly is valid also when t tends to ∞ , instead of to a $t^0 \neq \infty$. But if $J'' \rightarrow +\infty$ as $t \rightarrow t^0$, then, if $t^0 = \infty$, two quadratures show that $J \rightarrow +\infty$; so that the assumption $J \rightarrow 0$ leads to a contradiction.

Heliocentric Coordinates

§340. Since $\sum m_i \xi_i = 0$ is an identity in a barycentric coordinate system ξ , the problem with $3n$ degrees of freedom, which concerns the $3n$ coordinate vectors $\xi_i = \xi_i(t)$, can be reduced to one with $3(n-1)$ degrees of freedom, which concerns only $n-1$ of the n vectors ξ_i , say the vectors ξ_1, \dots, ξ_{n-1} ; or, what is the same thing, the $n-1$ differences

$$(1) \quad x_i = \xi_i - \xi_n, \quad (i = 1, \dots, n-1).$$

In fact, if the $n-1$ vectors (1) are known, ξ_n follows from $\sum m_i \xi_i = 0$, and ξ_1, \dots, ξ_{n-1} then follow from (1). The position vectors (1) of m_1, \dots, m_{n-1} with reference to m_n will be called heliocentric coordinates, m_n being defined to be "Sun" (even when m_n is not the largest of the n masses). Accordingly, a heliocentric coordinate system x is one having its origin at m_n and possessing coordinate axes which are parallel to those of an inertial coordinate system ξ at every t .

Since it is easily verified from (5), §314 and from the criterion (14), §318, that a heliocentric coordinate system is not, in general, inertial, the Lagrangian equations in terms of the heliocentric coordinates x_i

cannot be obtained simply by writing x for ξ in (5), §314, and must, therefore, be calculated in detail. Furthermore, this calculation cannot be based directly on the rule of §95. In fact, this rule assumes that the transformation formulae have a non-vanishing Jacobian and represent, therefore, a veritable transformation. But this condition is not satisfied by the heliocentric transformation, since (1) replaces the n coordinate vectors ξ_1, \dots, ξ_n by only $n - 1$ of their linear combinations.

§341. In order to avoid this difficulty, adjoin to (1), for a moment, the n -th linear combination,

$$(2) \quad x_n = \mu^{-1} \sum m_i \xi_i, \quad (\mu = \sum m_i),$$

of the n inertial coordinate vectors ξ_i (which need not be barycentric); so that the coordinate vector (2) of the centre of mass is considered as the n -th variable, instead of as the constant $x_n = 0$. It is easily seen that the conservative linear transformation (1)–(2) of ξ_1, \dots, ξ_n into x_1, \dots, x_n has the unique inverse

$$(3) \quad \xi_j = x_j - \mu^{-1} \sum_{i=1}^{n-1} m_i x_i + x_n, \quad \xi_n = -\mu^{-1} \sum_{i=1}^{n-1} m_i x_i + x_n, \\ (j = 1, \dots, n - 1),$$

and, correspondingly, a non-vanishing determinant ($= 1$). Hence, the rule of §95 is applicable to the transformation (1)–(2).

In the following explicit application of this rule, it will be convenient to use the summation symbols \sum^0, \sum^* which result by writing $n - 1$ for n in (4₁)–(4₂), §314; so that

$$(4_1) \quad \sum = \sum_{j=1}^n, \quad \sum^0 = \sum_{j=1}^{n-1};$$

$$(4_2) \quad \sum^* = \sum_{1 \leq j < k \leq n}, \quad \sum^+ = \sum_{1 \leq j < k \leq n-1} \quad (\sum^* = \sum^+ + \sum^0).$$

First, (3) implies that if $j < k$, then $\xi_j - \xi_k = x_j - x_k$ or $\xi_j - \xi_k = x_j$ according as $k = 1, \dots, n - 1$ or $k = n$. Hence, $U \equiv \sum^* m_j m_k / \rho_{jk}$, where $\rho_{jk} = |\xi_j - \xi_k|$, is transformed by (3) into

$$(5) \quad U = \sum^+ m_j m_k / \rho_{jk} + m_n \sum^0 m_j / \rho_{jn}, \quad \text{where} \\ \rho_{jk} = |x_j - x_k|, \quad \rho_{jn} = |x_j|.$$

On the other hand, substitution of (3) into $T \equiv \frac{1}{2} \sum m_i \xi_i'^2$ readily gives

$$(6_1) \quad T = \bar{T} + \frac{1}{2}\mu x_n'^2; \quad (6_2) \quad \bar{T} = \frac{1}{2} \sum^0 m_i x_i'^2 - \frac{1}{2} (\sum^0 m_i x_i')^2 / \mu;$$

$$(6_3) \quad \mu = \sum m_i.$$

Consequently, the Lagrangian function $L(\xi', \xi) \equiv T + U$ is transformed by (3) into $\bar{T} + \frac{1}{2}\mu x_n'^2 + U$, where U is given by (5) and contains, in view of (4₁)–(4₂), only the $n - 1$ heliocentric vectors (1); so that x_n is an ignorable coordinate in the sense of §182. Furthermore, \bar{T} is, in view of (6₂) and (4₁), free of x_n' . Consequently, the transformed Lagrangian function is of the form $\bar{L} + \frac{1}{2}\mu x_n'^2$, where $\bar{L} = \bar{T} + U$ does not contain x_n' , x_n . Hence, the Lagrangian equations $[\bar{L} + \frac{1}{2}\mu x_n'^2]_{x_i} = 0$, where $i = 1, \dots, n$, split into the system

$$(7) \quad [\bar{L}]_{x_i} \equiv [\bar{T} + U]_{x_i} = 0, \text{ i.e., } (\bar{T}_{x_i'})' = U_{x_i}; \quad i = 1, \dots, n - 1,$$

which is free of x_n' , x_n , and into the equation $[\frac{1}{2}\mu x_n'^2]_{x_n} \equiv \mu x_n'' = 0$, which is, in view of (2), equivalent to §317 bis.

Accordingly, x_n is a linear function of t , and the integration constants contained in $x_n = x_n(t)$ can, by §322, be chosen to be 0 without loss of generality. Then

$$(8) \quad x_n = 0$$

for every t , which means, by (2), that the inertial coordinates ξ_1, \dots, ξ_n occurring in (1) become barycentric. Finally, (7) is a Lagrangian system with $3(n - 1)$ degrees of freedom and contains, as desired, the heliocentric coordinates (1) only.

Needless to say, (7) does not possess the six integrals which correspond to those found in §317. In fact, these integrals are already expressed by (8) and have been used precisely in reducing the degree of freedom from $3n$ to $3(n - 1)$. On the other hand, (7) has the $3 + 1$ integrals which represent the three integrals found in §316 and the energy integral.

These integrals are

$$(9_1) \quad \sum^0 m_i x_i \times x_i' - (\sum^0 m_i x_i) \times (\sum^0 m_i x_i') / \mu = C;$$

$$(9_2) \quad \bar{T} - U = h,$$

where C, h are the same barycentric integration constants as in §322. In fact, substitution of (3) into $\sum m_i \xi_i \times \xi_i' = C$ easily leads to (9₁), if use is made of (4₁), (6₃) and (8). Similarly, (9₂) follows from (6₁), (6₂) and (8), since $T - U = h$.

An equivalent formulation of (9₂) is

$$(10_1) \quad J'' = 2U + 4h; \quad (10_2) \quad J = \sum^0 m_i x_i^2 - (\sum^0 m_i x_i)^2 / \mu.$$

In fact, the representation (10₂) of $J = \sum m_i \xi_i^2$ is clear from the proof of the representation (6₁)–(6₂) of $T = \frac{1}{2} \sum m_i \xi_i'^2$ and from (8); while (10₁) is the same thing as (2₄), §322.

§342. In order to obtain the explicit form of the heliocentric equations of motion, all that one has to do is to carry out the differentiations assigned by the Lagrangian form (7) of these equations, and then solve the resulting relations with respect to the heliocentric accelerations x_1'', \dots, x_{n-1}'' . It will be shown that the resulting explicit form of (7) can be written as

$$(11_1) \quad x_i'' + (m_n + m_i) \frac{x_i}{|x_i|^3} = {}^i\Omega_{x_i}; \quad i = 1, \dots, n-1;$$

$$(11_2) \quad {}^i\Omega = \sum_{k=1}^{n-1} m_k \left(\frac{1}{|x_k - x_i|} - \frac{x_k \cdot x_i}{|x_k|^3} \right),$$

${}^i\Omega_{x_i}$ in (11₁) denoting the x_i -gradient of the scalar sum (11₂), in which the dash of \sum means that $k \neq i$.

First, substitution of (6₂) into (7) gives

$$m_i x_i'' - m_i \sum^0 m_j x_j'' / \mu = U_{x_i}; \text{ hence, } m_n \mu^{-1} \sum^0 m_j x_j'' = \sum^0 U_{x_j}$$

follows by summation, since $1 - \sum^0 m_i / \mu = m_n / \mu$, by (6₃) and (4₁). On the other hand, it is seen from (5), (4₁)–(4₂) and from the meaning ($k \neq i$) of the summation symbol $\sum_{k=1}^{n-1}$, that

$$U_{x_i} = \sum_{k=1}^{n-1} m_i m_k \frac{x_k - x_i}{|x_k - x_i|^3} - m_n m_i \frac{x_i}{|x_i|^3}; \text{ hence,}$$

$$\sum^0 U_{x_j} = -m_n \sum^0 m_j \frac{x_j}{|x_j|^3},$$

the double sum omitted in $\sum^0 U_{x_j}$ being 0 in view of its skew-symmetry. The four relations contained in the last three formula lines clearly imply that

$$(12) \quad x_i'' + (m_n + m_i) \frac{x_i}{|x_i|^3} = \sum_{k=1}^{n-1} m_k \left(\frac{x_k - x_i}{|x_k - x_i|^3} - \frac{x_k}{|x_k|^3} \right);$$

$$i = 1, \dots, n-1, (k \neq i).$$

Finally, comparison of (12) with the definition (11₂) completes the proof of (11₁).

§343. Let $n = 2$. Then there exists only $n - 1 = 1$ vector equation (12), and the summation at the right of (12) is vacuous. Consequently, if x denotes the single $x_i = x_1$, one can write (1) and (12) as

$$(13) \quad x = \xi_1 - \xi_2 \text{ and} \\ x'' = - (m_2 + m_1) \frac{x}{|x|^3} \equiv V_x, \text{ where } V = \frac{m_1 + m_2}{|x|} \equiv \frac{\mu}{|x|}.$$

Hence, if $\mu = \sum m_i$ is chosen as the unit of mass, then $V = |x|^{-1}$; so that $x'' = V_x$ is, in view of §207, precisely the problem treated in §241–§312.

Accordingly, the content of the consequence (13) of (12) is that the problem of $n = 2$ bodies can be reduced to the problem of a single body moving in a static field of gravitation which has radial symmetry and is generated by a hypothetical body μ ; the latter body having the position of the "Sun" $m_n = m_2$, and a mass represented by the joint mass $m_1 + m_2$ of the "planet" m_1 and of m_2 .

The passage from the attracting mass m_2 to the greater mass $\mu = m_1 + m_2$ introduces, of course, a change in the original Lagrangian equations, since it introduces an additional force. The appearance of such a force is sufficiently explained by §340.

Needless to say, (13) remains valid on interchanging the subscripts 1, 2; so that either of the masses m_1, m_2 can be chosen as "Sun."

Since $\mu = m_1 + m_2$, one has $m_1 - m_1^2/\mu = m_1 m_2/\mu$; so that, by (5), (6₂), the integrals (9₂) and (9₁) of (13) become

$$(14_1) \quad \frac{1}{2}x'^2 - \frac{m_1 + m_2}{|x|} = \frac{m_1 + m_2}{m_1 m_2} h; \quad (14_2) \quad x \times x' = \frac{m_1 + m_2}{m_1 m_2} C.$$

§343 bis. If $n = 3$, then (1), (12) can be written as

$$(15_1) \quad x_1 = \xi_1 - \xi_3 \quad x_2 = \xi_2 - \xi_3;$$

$$(15_2) \quad x_1'' = q_{11}x_1 + q_{12}x_2, \quad x_2'' = q_{21}x_1 + q_{22}x_2,$$

where the scalars q are abbreviations for the four combinations

$$(16) \quad q_{\alpha\alpha} = - \frac{m_\alpha + m_3}{|x_\alpha|^3} - \frac{m_\beta}{|x_1 - x_2|^3}, \\ q_{\alpha\beta} = \frac{m_\beta}{|x_1 - x_2|^3} - \frac{m_\beta}{|x_\beta|^3}, \quad (\alpha \neq \beta = 1, 2),$$

of the two unknown vector functions x_1, x_2 of t (so that (15₂) is a non-linear system, of course).

§344. One has to expect a simplification of the problem (15₂) of $n = 3$ bodies if the two 3-vector differential equations (15₂) go over into each other by interchanging the subscripts 1 and 2. This will be the case if and only if

$$(17) \quad q_{11} \equiv q_{22} \quad \text{and} \quad q_{12} \equiv q_{21}$$

where the identity sign refers to t .

If, in addition, $x_1 \times x_2 \equiv 0$, so that the two vectors x_1, x_2 are collinear for every t , then (15₁) shows that the solution under consideration is collinear in the sense of §329, while (17) and (16) imply that $|x_1| \equiv |x_2|$, i.e., that m_3 lies at the mid-point of the segment m_1m_2 for every t . If, on the other hand, $q_{12} \equiv 0$ in (17), then (16) shows that $|x_1| \equiv |x_1 - x_2| \equiv |x_2|$, which means, by (15₁), that the three points ξ_1, ξ_2, ξ_3 occupied by the three masses form an equilateral triangle for every t . Hence, if the symmetry condition (17) is satisfied in the general equations (15₂), and if, in addition, either $x_1 \times x_2 \equiv 0$ or $q_{12} \equiv 0$, then the three bodies move so as to form a configuration which remains homographic to itself when t varies. Since later on (§369–§382) a general theory of all homographic solutions will be developed, the investigation of the case (17) may be restricted by the assumptions

$$(18_1) \quad x_1 \times x_2 \neq 0; \quad (18_2) \quad q_{12} \neq 0.$$

The object of the following considerations is an enumeration of all those solutions of the problem of $n = 3$ bodies for which the symmetry assumption (17) is compatible with (18₁)–(18₂).

Let a solution of the problem of $n = 3$ bodies be called an isosceles solution if the three bodies form, for every t , an isosceles triangle which can change its position and size when t varies and which is, in addition, such as to be neither a degenerate triangle (i.e., a segment) for every t nor an equilateral triangle for every t . Then the enumeration problem mentioned above requires the enumeration of all isosceles solutions belonging to the case of two equal masses on the base. In fact, (16) shows that the assumption (17) is equivalent to the pair of conditions

$$(19_1) \quad m_1 = m_2; \quad (19_2) \quad |x_1| \equiv |x_2|,$$

where $|x_1|$, $|x_2|$ are, in view of (15₁), the lengths of the sides of the triangle which are opposite the two equal masses.

Actually, it turns out (§389) that an isosceles solution is possible only when the two masses on the base are equal.

§345. It will be convenient to replace the pair of vectors (15₁) by their linear combinations $\frac{1}{2}(x_1 \pm x_2)$; so that

$$(20_1) \quad X_1 = \frac{1}{2}(x_1 + x_2), \quad X_2 = \frac{1}{2}(x_1 - x_2);$$

$$(20_2) \quad X_1' = (q_{11} + q_{12})X_1, \quad X_2' = (q_{11} - q_{12})X_2,$$

by (15₂) and (17). Then

$$(21_1) \quad X_1 \cdot X_2 = 0; \quad (21_2) \quad X_1 \cdot X_2' = -X_2 \cdot X_1';$$

$$(21_3) \quad X_1 \cdot X_2'' = 0, \quad X_2 \cdot X_1'' = 0.$$

In fact, (15₁) shows that the coordinate vectors of m_1 , m_2 , m_3 in the heliocentric coordinate system x (with m_3 as Sun) are x_1 , x_2 , 0 , respectively. Hence, $\frac{1}{2}(x_1 + x_2)$ is the mid-point of the base m_1m_2 of the isosceles triangle, while the vector $x_1 - x_2$ is perpendicular to this base; so that the vectors (20₁) are perpendicular. This proves (21₁), while (21₂) follows by differentiation of (21₁); and (21₃) is clear from (21₁) and (20₂).

Furthermore, there exist two constant vectors A_1 , A_2 such that, for every t ,

$$(22_1) \quad X_1 \times X_1' = A_1, \quad X_2 \times X_2' = A_2;$$

$$(22_2) \quad A_1 \cdot X_1 = 0, \quad A_2 \cdot X_2 = 0; \quad (22_3) \quad (X_1 \cdot X_2')^2 = A_1 \cdot A_2.$$

In fact, (20₂) implies, for $i = 1, 2$, that $X_i \times X_i'' \equiv 0$, i.e., that $X_i \times X_i' = \text{const.}$ This proves not only (22₁) but also (22₂), since $(Y \times Z) \cdot Y \equiv 0$. Finally, on applying the identity (3), §65 to $a = X_1$, $b = X_1'$, $c = X_2$, $d = X_2'$ and then using (21₁), (21₂) and (22₁), one obtains (22₃).

Differentiation of the constant (22₃) gives $X_1' \cdot X_2' = -X_1 \cdot X_2''$, which means, by (21₃), that $X_1' \cdot X_2'$ vanishes identically. Hence, the same holds for the derivative $(X_1' \cdot X_2')' \equiv X_1'' \cdot X_2' + X_2'' \cdot X_1'$. Substituting X_1'' , X_2'' from (20₂) and then using (21₂), one sees therefore that $2q_{12}X_1 \cdot X_2'$ vanishes identically, which means, by (18₂) and (22₃), that $A_1 \cdot A_2 = 0$. In other words, A_1 , A_2 is a perpendicular pair of vectors. Since (21₁) and (22₂) show that the same holds for any of the three pairs X_1 , X_2 ; A_1 , X_1 ; A_2 , X_2 , it follows that the four vectors A_1 , A_2 , X_1 , X_2 are mutually perpendicular. Consequently,

at least one of these four 3-vectors must vanish. But neither X_1 nor X_2 can vanish identically. In fact, the assumption (18₁) is, in view of (20₁), equivalent to $X_1 \times X_2 \neq 0$ and implies, therefore, that $X_1 \neq 0$ and $X_2 \neq 0$.

§346. It follows that

(I) at least one of the two constant vectors A_1, A_2 vanishes;

(II) save at most for isolated values of t , the vectors $X_1 = X_1(t)$, $X_2 = X_2(t)$ do not vanish and are perpendicular, so that, in particular, they determine a plane through the origin of the X -space;

(III) if $A_\alpha \neq 0$, the constant direction of A_α is perpendicular to the plane mentioned under (II);

(IV) if $A_\alpha = 0$, then, as seen from (22₁), the vector $X_\alpha = X_\alpha(t)$ moves on a line which goes through the origin of the X -space and does not vary with t ;

(V) in all cases, the plane mentioned under (II) does not vary with t (this is implied by (III) or by a two-fold application of (IV) according as the assumption $A_\alpha = 0$ of (IV) is not or is satisfied for both values of α).

The representation of X_1, X_2 in terms of ξ_1, ξ_2, ξ_3 is

$$(23_1) \quad X_1 = \nu \xi_3, \quad X_2 = \frac{1}{2}(\xi_1 - \xi_2);$$

$$(23_2) \quad \nu = -\mu/(\mu - m_3), \quad (\mu = \sum m_i).$$

This is clear from (19₁) and from the barycentric assumption $\sum m_i \xi_i = 0$, since $X_1 = \frac{1}{2}(\xi_1 + \xi_2) - \xi_3$, $X_2 = \frac{1}{2}(\xi_1 - \xi_2)$, by (20₁) and (15₁).

Notice that (I) allows three cases only:

$$(i) \quad A_1 \neq 0 = A_2; \quad (ii) \quad A_1 = 0 \neq A_2; \quad (iii) \quad A_1 = 0 = A_2;$$

the case (iii) being that in which (IV) is applicable twice and (III) fails, while (III) and (IV) are applicable exactly once in the cases

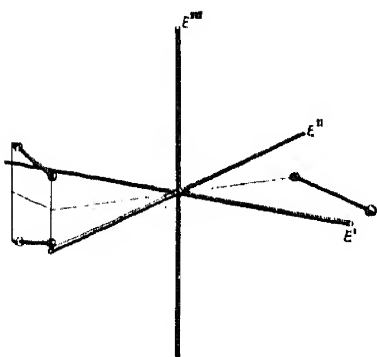


FIG. 12_i

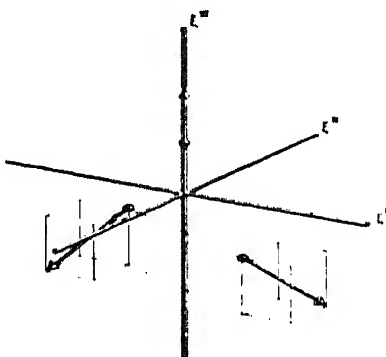


FIG. 12_{ii}

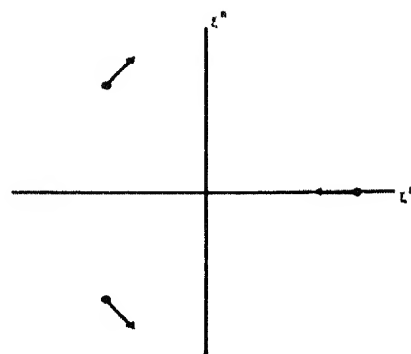


FIG. 12_{iii}

(i) and (ii), finally (II) and (V) hold in all three cases. Hence, the constant (23_2) being a non-vanishing scalar, it is seen from (23_1) that every isosceles motion of the two equal masses and of m_3 must present, for every t , one of three symmetric situations.

These are indicated in the three figures which correspond to the cases (i), (ii), (iii), and assume that the position of the barycentric coordinate system has been chosen suitably. The arrows indicate velocity vectors. The velocity vector of m_3 lies in the (ξ^I, ξ^{II}) -plane of the first figure. The masses m_1 , m_2 and their velocity vectors possess symmetry with respect to the ξ^{III} -axis in the second figure. In the third figure, the plane of the paper is the (ξ^I, ξ^{II}) -plane.

§346 bis. According to the figures, only the case (iii) is planar in the sense of §324; while the difference between the two non-planar cases is that the motion has a fixed plane of symmetry (the (ξ^I, ξ^{II}) -plane) in the case (i), but a fixed axis of symmetry (the ξ^{III} -axis) in the case (ii). In the case (iii), there is a fixed axis of symmetry within the plane of movement.

It is easily shown that in the cases (i) and (iii) there must occur a collision of m_1 and m_2 , and that this is impossible in the case (ii).

Finally, comparison of the figures with (5), §322 shows that, for reasons of symmetry, (6), §323 is satisfied in all three cases, with $C = 0$ in the case (iii); while §326 implies that $C \neq 0$ in the cases (i) and (ii).

It follows that the condition $C = 0$ of §335 is only a necessary, but not a sufficient, condition for a simultaneous collision of all bodies, since, in the case (iii), such a collision may be excluded by a suitable choice of the remaining integration constants.

§347. Since $m_1 = m_2$, it is clear for reasons of symmetry that any of the symmetry conditions (i), (ii), (iii) for the six 3-vectors $\xi_i(t)$, $\xi'_i(t)$ is satisfied for every t , if it is satisfied for an initial $t = t^0$ by the choice of the integration constants $\xi_i(t^0)$, $\xi'_i(t^0)$. Thus, it is evident that each of these three kinds of isosceles solutions actually exists. But the object of §345–§346 was to prove that, on the assumption (19_1) , these obvious types of isosceles solutions exhaust the totality of all isosceles solutions (cf. also the last remark of §344). As will be seen at the end of §374 bis, this fact is by no means evident.

§347 bis. It may be mentioned that the general problem (15_2) with $3(n - 1) = 3 \cdot 2 = 6$ degrees of freedom immediately reduces, in any of the three isosceles cases, to a conservative dynamical sys-

tem which has 2 degrees of freedom but no known integral, except for the energy integral; so that none of the three cases (i), (ii), (iii) can be integrated explicitly.

This completes the discussion of the particular problem which arose in §344. In what follows, quite a different application of the general heliocentric Lagrangian equations (12) will be considered.

Binary Collisions

§348. Suppose that a given solution of the problem of n bodies is known to be such that, at every date of a certain t -interval,* the distance between either of two bodies, say m_1 and m_n , and any of the $n - 2$ remaining bodies, m_2, \dots, m_{n-1} , exceeds a fixed positive lower bound; so that, by (1),

$$(24) \quad |x_1 - x_k| > \text{const.} > 0 \quad \text{and} \quad |x_k| > \text{const.} > 0 \\ \text{for } k = 2, \dots, n - 1.$$

Then it is natural to ask how to estimate, during the time interval under consideration, the deviation of the actual motion of m_1 and m_n from the motion of m_1 and m_n which would arise if m_2, \dots, m_{n-1} did not exist; in other words, how to estimate the error which one commits by writing $x_n (= \xi_1 - \xi_n)$ and m_n for x and m_2 in the equations (13)–(14₂) of the problem of two bodies.

Such estimates can easily be obtained by observing that (12) reduces, if $i = 1$, to

$$(25_1) \quad x_1'' + (m_n + m_1) \frac{x_1}{|x_1|^3} = f;$$

$$(25_2) \quad f = \sum_{k=2}^{n-1} m_k \left(\frac{x_k - x_1}{|x_k - x_1|^3} - \frac{x_k}{|x_k|^3} \right).$$

As an application of (25₁)–(25₂), it will be shown that

$$(26_1) \quad \left| \left(\frac{x_1}{|x_1|} \right)' \right| \leq \frac{|x_1 \times x_1'|}{x_1^2}; \quad (26_2) \quad \frac{|(x_1 \times x_1')'|}{x_1^2} < \text{Const.},$$

where the constant depends only on the constants occurring in the assumption (24) and on the given values of the masses m_i .

First, $|a|^2 b - (a \cdot b)a = (a \times b) \times a$ for arbitrary 3-vectors a, b . Hence, on placing $a = x_1$, $b = x_1'$ and noting that $x_1 \cdot x_1' = |x_1| |x_1'|$,

* This interval may be finite or infinite.

one obtains $|x_1| \{ |x_1| x'_1 - |x_1|' x_1 \} = (x_1 \times x'_1) \times x_1$. Since $x_1 \neq 0$ and $|c \times d| \leq |c| |d|$, it follows that $|\{ |x_1| x'_1 - |x_1|' x_1 \}| \leq |x_1 \times x'_1|$. This proves (26₁).

Next, $x_1 \times x_1'' \equiv (x_1 \times x'_1)'$ and $x_1 \times x_1 \equiv 0$; so that vector multiplication of (25₁)–(25₂) by x_1 shows that

$$(x_1 \times x'_1)' = \sum_{k=2}^{n-1} m_k (\beta_k^3 - \alpha_k^3) x_k \times x_1, \quad \text{where}$$

$$\alpha_k = |x_k - x_1|^{-1}, \quad \beta_k = |x_k|^{-1}.$$

Hence, an application of (24) and of the identity $\beta^3 - \alpha^3 = (\beta - \alpha)(\alpha^2 + \alpha\beta + \beta^2)$ gives

$$|(x_1 \times x'_1)'| \leq \text{Const.} \sum_{k=2}^{n-1} ||x_k - x_1|^{-1} - |x_k|^{-1}| |x_1 \times x_k|.$$

Since $|x_1 \times x_k| \leq |x_1| |x_k|$, while

$$\begin{aligned} & ||x_k - x_1|^{-1} - |x_k|^{-1}| \\ & \leq \frac{||x_k| - |x_k - x_1||}{|x_k - x_1| |x_k|} \leq \frac{|x_1|}{|x_k - x_1| |x_k|} < \text{Const.} \frac{|x_1|}{|x_k|}, \text{ by (24),} \end{aligned}$$

it follows that (26₂) holds for some Const.

It should be mentioned for later reference that scalar multiplication of (25₁) by $\frac{1}{2}x'_1$ and x_1 gives

$$(27_1) \quad g' = \frac{1}{2}f \cdot x'_1, \quad \text{where} \quad g \equiv \frac{1}{2}x_1'^2 - \frac{m_n + m_1}{|x_1|};$$

$$(27_2) \quad (\tfrac{1}{2}x_1^2)'' - x_1'^2 + \frac{m_n + m_1}{|x_1|} = f \cdot x_1,$$

respectively, since $(\frac{1}{2}x_1'^2)' = x_1' \cdot x_1''$ and $(\frac{1}{2}x_1^2)'' = x_1 \cdot x_1'' + x_1^2$.

§349. A given solution of the problem of n bodies will be said to lead, at a date $t = t^0 (\neq \infty)$, to a binary collision if, on the one hand, the distance between two of the n bodies m_i , say between m_1 and m_n , tends, as $t \rightarrow t^0$, to 0, and, on the other hand, any of the remaining $\frac{1}{2}n(n-1) - 1$ mutual distances ultimately* surpasses a fixed positive lower bound. It is clear that the influence of the $n - 2$ bodies

* This word is meant in the sense defined by the footnote to §335.

m_2, \dots, m_{n-1} on \tilde{m}_1 and m_n , when compared with the action of m_1 and m_n on each other, will ultimately become quite unimportant. Hence, it can be expected that the position vector $x_1(t) = \xi_1(t) - \xi_n(t)$ of m_1 with reference to m_n (cf. (1), §340) behaves, at the date $t = t^0$ of the binary collision, in about the same way as if the collision of m_1 and m_n were to take place merely in a problem of two bodies (cf. §343 and §268). But in the problem of two bodies, a collision is possible only in case of rectilinear motion; furthermore, the energy integral implies that the relative speed must become infinite in case of a collision in the problem of two bodies; finally, (10₁)–(10₂) and (5), when applied to the case of a collision in the problem of only two bodies, imply a relation between the mutual distance of two bodies and its second derivative.

It turns out that the corresponding facts hold in the limit as $t \rightarrow t^0$, if there is a binary collision of m_1 and m_n at $t = t^0$. This will be proved by showing that, no matter what are the $n - 2$ masses m_2, \dots, m_{n-1} which do not participate in the binary collision, for the position vector of m_1 with reference to m_n one has

$$(28_1) \quad x_1 \times x_1' \rightarrow 0; \quad (28_2) \quad (\tfrac{1}{2}x_1^2)'' |x_1| \rightarrow m_n + m_1;$$

$$(28_3) \quad \tfrac{1}{2}x_1'^2 |x_1| \rightarrow m_n + m_1,$$

when $t \rightarrow t^0$, $x_1 \equiv \xi_1 - \xi_n \rightarrow 0$.

While the paths of m_1 and m_n are not, in general, plane curves, (28₁) shows that these paths are practically rectilinear when t is close to t^0 .

§349 bis. In order to prove (28₁)–(28₃), notice first that $|f| < \text{const.}$, by (25₂) and (24). Hence, on multiplying (27₂) by $|x_1|$ and then letting $|x_1| \rightarrow 0$, one obtains

$$|x_1| (\tfrac{1}{2}x_1^2)'' - |x_1| x_1'^2 + m_n + m_1 \rightarrow 0, \quad (t \rightarrow t^0);$$

a relation which shows that (28₂) is equivalent to (28₃). On the other hand, (28₃) implies (28₁). In fact, it is clear from (28₃), where $x_1 \rightarrow 0$, that $|x_1'| \rightarrow +\infty$, hence $|x_1| |x_1'| \rightarrow 0$; so that (28₁) follows from $|x_1 \times x_1'| \leq |x_1| |x_1'|$. Thus, it will be sufficient to prove (28₃).

It is assumed (cf. §348) that, while $|\xi_1 - \xi_n| = |x_1|$ tends to 0, none of the remaining $\frac{1}{2}n(n-1) - 1$ mutual distances $|\xi_j - \xi_k| = \rho_{jk}$ comes arbitrarily close to 0, as $t \rightarrow t^0$. Since the force function U is a linear combination of the reciprocal values of all $\frac{1}{2}n(n-1)$

distances, it follows from the energy integral $T - U = h$ that the kinetic energy T is ultimately* less than $\text{const.}/|x_1|$. But the formulae of §341 show that T is a positive definite quadratic form in the components of the velocities x'_1, \dots, x'_{n-1} ; so that the single speed $|x'_1|$ clearly is majorized by $\text{Const. } T^{\frac{1}{2}}$. Consequently, $|x'_1| < \text{Const.}/|x_1|^{\frac{1}{2}}$.

On the other hand, on expanding (25₂) according to powers of $|x_1|$ and of the components of the 3-vector x_1 , one sees from (24) and the assumption $x_1 \rightarrow 0$ that, ultimately, $|f| < \text{const.}|x_1|$. Accordingly, $|f||x'_1| < \text{Const.}|x_1|^{\frac{1}{2}}$. Since $|f \cdot x'_1| \leq |f||x'_1|$ and $|x_1| \rightarrow 0$, it follows that $f \cdot x'_1 \rightarrow 0$, as $t \rightarrow t^0$. Hence, (27₁) shows that the derivative of the difference $g = g(t)$ which is defined by (27₁) tends, as $t \rightarrow t^0$, to 0. In view of an elementary rule in calculus,† this is possible only when the function $g = g(t)$ itself tends to a finite limit, as $t \rightarrow t^0$. Since $x_1 \equiv x_1(t) \rightarrow 0$, it follows that $|x_1|g \rightarrow 0$. This, when compared with the definition (27₁) of the difference g , completes the proof of (28₃).

§350. The distance ρ_{1n} between the two bodies m_1, m_n which participate, as $t \rightarrow t^0$, in the binary collision is $|\xi_1 - \xi_n| = |x_1|$, by (1). Hence, the relation (28₂), proved in §349 bis, may be written in the form $\lim \rho_{1n}(\rho_{1n}^2)'' = 2(m_n + m_1)$. Clearly, this formula for a binary collision goes over into the formula (18₁), §337 for a simultaneous collision of all n bodies, if one lets correspond $\rho_{1n}^2(t) = \rho_{1n}^2 \equiv (\xi_1 - \xi_n)^2$ to $J(t) = J \equiv \sum m_i \xi_i^2$, and $2(m_n + m_1)$ to the positive constant μ_0 . Since (18₁), §337 was shown to imply the asymptotic relations (18₂) and (17₁)–(17₂) of §336–§337, it follows that these relations remain valid if one replaces J by ρ_{1n}^2 and μ_0 by $2(m_n + m_1)$. Consequently, the distance $\rho_{1n} = \rho_{1n}(t)$ behaves, in case of a binary collision at $t = t^0$, in such a way that

$$(29) \quad \rho_{1n} \sim \left[\frac{9}{2}(m_1 + m_n) \right]^{\frac{1}{2}} (t - t^0)^{\frac{2}{3}} \quad \text{as } t \rightarrow t^0;$$

furthermore, (29) remains valid on two-fold differentiation with respect to t . Since $\rho_{1n} = |x_1|$, one sees from (29) and (28₃) that the relative speed $|x'_1| \equiv |\xi'_1 - \xi'_n|$ of the colliding bodies becomes infinite in the order $(t - t^0)^{-\frac{1}{3}}$, as $t \rightarrow t^0$.

§351. The relation (28₁), where $x_1 = \xi_1 - \xi_n$, states merely that

* Cf. the preceding footnote.

† Cf. the last remark of §351 below.

the paths of the colliding bodies m_1, m_n become parallel to each other as $t \rightarrow t^0$; it does not state that these paths touch each other at $t = t^0$. In fact, a collision of m_1 and m_n is defined by the condition $|\xi_1 - \xi_n| \rightarrow 0$, which, in itself, would allow that m_1 and m_n move, before colliding, along spirals without asymptotes in such a way that the directions of the tangents of the two paths $\xi = \xi_1(t), \xi = \xi_n(t)$ do not tend to limiting positions as $t \rightarrow t^0$. This possibility will now be excluded by showing that the binary collision of m_1 and m_n must take place at a definite angle. More precisely, the unit vector $x_1/|x_1|$, where $x_1 = \xi_1 - \xi_n$, tends to a limiting position and, in addition, the direction of $x_1/|x_1|$ ultimately varies so slowly that its derivative $(x_1/|x_1|)'$ tends to 0 as $t \rightarrow t^0$.

First, on integrating the derivative of $x_1 \times x_1'$ between two dates t, t^* , then keeping t fixed in some position close to t^0 , but varying t^* by letting $t^* \rightarrow t^0$, one sees from (28₁) that

$$x_1(t) \times x_1'(t) = \int_{t^0}^t (x_1 \times x_1')' d\bar{t};$$

hence,

$$|x_1(t) \times x_1'(t)| < \text{Const.} \left| \int_{t^0}^t x_1(\bar{t})^2 d\bar{t} \right|,$$

by (26₂). But (28₂) clearly implies that x_1^2 is ultimately decreasing (the reason being the same as in §335 for J). Hence, the last integral is majorized by $x_1(t)^2 |t - t^0|$, and so the value of $|x_1 \times x_1'|/x_1^2$ at the date t is less than $\text{Const.} |t - t^0|$. Thus, on letting $t \rightarrow t^0$, one sees from (26₁) that the proof of $(x_1/|x_1|)' \rightarrow 0$ is complete. Finally, the existence of $\lim x_1/|x_1|$ follows by applying to the function $f = x_1/|x_1|$ the following remark (which becomes evident if one integrates f'):

If the derivative $f'(t)$ of a function $f(t)$ remains bounded as $t \rightarrow t^0 (\neq \infty)$, then $f(t)$ tends, as $t \rightarrow t^0$, to a finite limit.

§352. That part of the definition (§349) of a binary collision which concerns the colliding bodies m_1, m_n requires only that $|\xi_1 - \xi_n| \rightarrow 0$ as $t \rightarrow t^0$. This states only that the respective positions $\xi = \xi_1(t), \xi = \xi_n(t)$ of m_1, m_n in the barycentric inertial coordinate system ξ tend to each other as $t \rightarrow t^0$; a condition which, in itself, would allow that neither $\xi_1(t)$ nor $\xi_n(t)$ tends to a (finite or infinite) limit. Actually, the condition imposed in §349 on the ultimate behavior of the paths of m_2, \dots, m_{n-1} insures that there exists a finite common limit $\lim \xi_1 = \lim \xi_n$ as $t \rightarrow t^0$; so that the collision of m_1 and m_n

must take place at a well-determined point of the barycentric inertial Cartesian space ξ .

In fact, since $|\xi_1 - \xi_n| \rightarrow 0$, and since the masses are positive constants, the existence of a finite common limit $\lim \xi_1 = \lim \xi_n$ is equivalent to the existence of a finite limit for the function $m_1\xi_1 + m_n\xi_n$. But, the inertial coordinate system ξ being barycentric, $m_1\xi_1 + m_n\xi_n$ is identical with the sum of the $n - 2$ terms $-m_l\xi_l$, where $l = 2, \dots, n - 1$. Thus, it is sufficient to prove the existence of finite limits for the $n - 2$ position vectors ξ_l . Accordingly, application of the last remark of §351 to $f = \xi_l$ shows that it is sufficient to prove the existence of finite limits for the $n - 2$ velocity vectors ξ'_l . Consequently, application of the last remark of §351 to $f = \xi'_l$ shows that it is sufficient to prove the boundedness of the $n - 2$ acceleration vectors ξ''_l , as $t \rightarrow t^0$. This requires merely that the forces of gravitation acting on m_l , where $l = 2, \dots, n - 1$, remain bounded, as $t \rightarrow t^0$. Now, the equations of motion (1₁)–(1₃), §322 show that this condition is satisfied, since it is assumed (§349) that, except in the case $|\xi_j - \xi_k| = |\xi_1 - \xi_n|$ of the colliding bodies m_1, m_n , the distance $|\xi_j - \xi_k|$ between m_j and m_k , where $1 \leq j < k \leq n$, does not come arbitrarily close to 0, as $t \rightarrow t^0$.

Accordingly, ξ_l, ξ'_l tend, if $l \neq 1$ and $l \neq n$, to finite limits, say ξ_l^0, ξ'^0_l ; and, in addition, ξ_1, ξ_n tend to a common finite limit $\xi_1^0 = \xi_n^0$ which is distinct from all ξ_l^0 . On the other hand, ξ'_1, ξ'_n cannot tend to finite limits, since $|\xi'_1 - \xi'_n| \rightarrow +\infty$, by (28₃).

Incidentally, the $n - 2$ acceleration vectors ξ''_l , where $l = 2, \dots, n - 1$, not only remain bounded but also tend to finite limits, as $t \rightarrow t^0$. This is now clear from (1₁)–(1₃), §322 and from the existence of all the finite limits $\lim \xi_i$, where $i = 1, \dots, n$ and either $\lim \xi_j \neq \lim \xi_k$ or $j = 1, k = n$.

§353. The above results will now be shown to imply that if a solution of the problem of $n = 3$ bodies is not planar in the sense of §324 and leads, as $t \rightarrow t^0$, to a binary collision of two of the three bodies, then the collision of these two bodies takes place at a point situated within the invariable plane; while the path of the body which does not participate in the collision has at its point $t = t^0$ the invariable plane as tangent plane.

Since it is assumed that the solution is not planar, and since for $n = 3$ every solution is flat, §326 ensures that $C \neq 0$; so that there exists an invariable plane $C \cdot \xi = 0$. The statement is that the place $\xi_1^0 = \xi_3^0$ of the collision of m_1 and m_3 , as well as the position and the

velocity vectors $\xi_2^0, \xi_2'^0$ of m_2 at $t = t^0$, lie in this plane (while $\xi_2'^0 \neq 0$), where $\xi_1^0, \xi_2'^0$ denote the finite limits whose existence was proved in §352.

To this end, it will be shown that $\xi_2^0 \times \xi_2'^0 = \nu C$, where ν is a non-vanishing scalar. This will clearly imply that $\xi_2'^0 \neq 0$ and (upon scalar multiplication by $\xi_2^0, \xi_2'^0$) that $\xi_2'^0, \xi_2^0$ satisfy the equation $0 = C \cdot \xi$ of the invariable plane. But then the same will hold for ξ_1^0, ξ_3^0 , since $\xi_1^0 = \xi_3^0$ is a scalar multiple of ξ_2^0 , the sum of all three $m_i \xi_i^0$ being 0 in a barycentric coordinate system ξ .

Accordingly, it will be sufficient to prove that $\xi_2^0 \times \xi_2'^0$ is of the form νC , where $\nu \neq 0$. Actually, the proof of this fact will be independent of the assumption $C \neq 0$ and will, therefore, imply that $\xi_2^0 \times \xi_2'^0 = 0$ holds if and only if $C = 0$.

Thus, if $n = 3$, the limiting velocity vector $\xi_2'^0$ of m_2 is or is not situated within the line joining the limiting position ξ_2^0 of m_2 with the place $\xi_1^0 = \xi_3^0$ of the binary collision of m_1 and m_3 , according as there does not or does exist an invariable plane. Notice that the non-existence of the invariable plane (i.e., $C = 0$) is sufficient but not necessary for a planar solution, if $n = 3$ (cf. §326 with the last remark of §324).

§354. In order to prove that $\xi_2^0 \times \xi_2'^0 = \nu C$, where $\nu \neq 0$, notice first that, the coordinate system ξ being barycentric, the sum of all three $m_i \xi_i$ vanishes identically, as does the sum of all three $m_i \xi_i'$. On calculating from this pair of linear relations the vector product of $m_2 \xi_2$ and $m_2 \xi_2'$, and dividing the result by $m_1 m_3$, one sees that

$$\mu_{21} \mu_{23} (\xi_2 \times \xi_2') = \mu_{13} (\xi_1 \times \xi_1') + \mu_{31} (\xi_3 \times \xi_3') + (\xi_1 \times \xi_3') + (\xi_3 \times \xi_1'),$$

where $\mu_{ik} = m_i/m_k$. On the other hand, from (1), §340, where $n = 3$,

$$x_1 \times x_1' = (\xi_1 \times \xi_1') + (\xi_3 \times \xi_3') - (\xi_1 \times \xi_3') - (\xi_3 \times \xi_1').$$

On adding these two relations, multiplying the result by $m_1 m_3$, then letting $t \rightarrow t^0$ and using (28₁), one obtains

$$m_2^2 (\xi_2^0 \times \xi_2'^0) = (m_1 + m_3) \lim_{t \rightarrow t^0} \{ m_1 (\xi_1 \times \xi_1') + m_3 (\xi_3 \times \xi_3') \}.$$

But the last limit is $C = m_2 (\xi_2^0 \times \xi_2'^0)$, since the sum of all three $m_i \xi_i \times \xi_i'$ is the constant angular momentum C . This completes

the proof of $\xi_2^0 \times \xi_2'^0 = \nu C$, supplying for ν the value (> 0) of the ratio of $m_1 + m_3$ and $m_2(m_1 + m_2 + m_3)$.

Central Configurations

§355. From here on to the beginning of §361, the time parameter t will be supposed to have a fixed value; so that only the barycentric positions ξ_i , and not also the velocities ξ_i' or accelerations ξ_i'' , of the n masses will be considered. The force acting on m_i at the fixed date t is nevertheless defined, since it is the vector $U_{\xi_i} = U_{\xi_i}(\xi_1, \dots, \xi_n)$, where $U = \sum^* m_j m_i / |\xi_j - \xi_i|$. Similarly, also $J = \sum m_i \xi_i^2$ and its gradients $J_{\xi_i} = 2m_i \xi_i$ are defined at the given position.

The n position vectors ξ_i of the n bodies m_i will be said to form a central configuration with respect to the n fixed positive constants m_i , if the force of gravitation acting on m_i at the moment of the given configuration is proportional to the mass m_i and to the barycentric position vector ξ_i ; i.e., if $U_{\xi_i} = \sigma m_i \xi_i$ holds for $i = 1, \dots, n$ and for some scalar σ which is independent of i . Actually, the value of σ is then uniquely determined. In fact, $\sum \xi_i \cdot U_{\xi_i} = \sigma \sum m_i \xi_i^2$, where $\sum \xi_i \cdot U_{\xi_i} = -U$, since U is homogeneous of degree -1 ; so that $\sigma = -U/J$.

This, when combined with $J_{\xi_i} = 2m_i \xi_i$, shows that the conditions $U_{\xi_i} = \sigma m_i \xi_i$ for a central configuration of the m_i may be written as

$$(1) \quad \begin{aligned} JU_{\xi_i} &= -\frac{1}{2}UJ_{\xi_i}, \quad \text{i.e.,} \quad (JU^2)_{\xi_i} = 0; \quad i = 1, \dots, n \\ (U &= \sum^* m_j m_k / |\xi_j - \xi_k|, \quad J = \sum m_k \xi_k^2). \end{aligned}$$

A central configuration will be called flat if its n points ξ_i are contained in a plane, where the case of a collinear central configuration is not excluded; so that $n \geq 4$ is a necessary condition for a non-flat configuration.

It is clear that the notion of a central configuration is independent of the orientation of the barycentric coordinate system ξ . Furthermore, it is clear, for reasons of homogeneity, that if ξ_1, \dots, ξ_n form a central configuration with respect to m_1, \dots, m_n , then so do $\bar{\xi}_1, \dots, \bar{\xi}_n$ whenever $\bar{\xi}_i = \beta \xi_i$ holds for some $\beta > 0$ and for every i . Correspondingly, two central configurations which belong to the same m_i will be considered as identical not only if they are congruent in the sense of Euclidean geometry but also if they go over into each other upon a suitable change of the unit of length.

§356. Let $\Gamma = (\gamma_{ik})$ denote the n -matrix defined by

$$(2) \quad \Gamma = (\gamma_{ik}): \gamma_{ik} = \frac{m_i}{|\xi_i - \xi_k|} \text{ if } i \neq k, \gamma_{ii} = \frac{U}{J} - \sum_{j=1}^n{}' \frac{m_j}{|\xi_i - \xi_j|^3},$$

where the dash ' means that $j \neq i$. Denoting by ξ_i^ν , where $\nu = \text{I, II, III}$, the components of the n three-vectors ξ_i , and by Ξ^ν the three n -vectors whose components are the ξ_i^ν , one easily verifies that the n three-vector conditions (1) may be written in the form of the three n -vector conditions $\Gamma \Xi^\nu = 0$. In particular, a necessary condition for a central configuration is that $\det \Gamma = 0$, i.e., $r \leq n - 1$, where r denotes the rank of Γ . Notice that if Γ is given and ν is fixed, the n -vector condition $\Gamma \Xi^\nu = 0$ for the ξ_i^ν , when combined with the barycentric condition $\sum m_i \xi_i^\nu = 0$, determines the mutual ratios of the n scalars ξ_i^ν uniquely only when $r = n - 1$.

The necessary condition $r \leq n - 1$ may also be expressed by saying that 0 must be a characteristic number of the n -matrix Γ . Actually, r is always precisely the multiplicity of the root 0 of the characteristic equation of Γ . In fact, the matrix (2), while not symmetric, becomes symmetric if one multiplies its k -th column by m_k . Nevertheless, the characterization of central configurations in terms of the matrix Γ is often unmanageable, and will not be used in what follows.

§357. A more convenient characterization of central configurations may be obtained by expressing (1) in terms of the $\frac{1}{2}n(n-1)$ mutual distances $\rho_{ik} = |\xi_i - \xi_k|$, where $1 \leq i < k \leq n$.

First, by §322 bis,

$$(3_1) \quad J = \mu^{-1} \sum^* m_j m_l \rho_{jl}^2, \text{ while } U = \sum^* m_j m_l / \rho_{jl}; \quad (3_2) \quad \mu = \sum m_j.$$

However, one cannot replace (1) by the $\frac{1}{2}n(n-1)$ conditions $(JU^2)_{\rho_{ik}} = 0$, unless the ρ_{ik} are geometrically independent; which they are not, in general. For instance, it is known from analytic geometry that six positive numbers $\rho_{ik} = \rho_{ki}$, where $1 \leq i < k \leq 4$, do or do not represent the mutual distances between four co-planar but not collinear points according as there is or is not satisfied the geometrical condition

$$(4) \quad R = 0, \text{ where } R \equiv R(\rho_{12}, \rho_{13}, \rho_{14}, \rho_{23}, \rho_{24}, \rho_{34}) = \det \Delta,$$

Δ denoting the symmetric 5-matrix in which the i -th element of the k -th column is, if $i \neq k$, the square of ρ_{ik} or 1 according as $i = 1, 2, 3, 4$ or $i = 5$, while all five diagonal elements are 0. This is only an illus-

tration of the general fact that $\frac{1}{2}n(n-1)$ positive numbers $\rho_{ik} = \rho_{ki}$ represent the mutual distances between n distinct points of the Euclidean 3-space if and only if there are satisfied $p(\geq 0)$ independent conditions

$$(5) \quad R_s = 0; \quad s = 1, \dots, p, \quad \text{where } R_s \equiv R_s(\rho_{12}, \rho_{13}, \dots, \rho_{n-1, n})$$

is a rational function of the $\frac{1}{2}n(n-1)$ variables ρ_{ik} . The number p of these functions is given by the first, second or third of the relations

$$(6_1) \quad p = \frac{1}{2}(n-1)(n-2); \quad (6_2) \quad p = \frac{1}{2}(n-2)(n-3);$$

$$(6_3) \quad p = \frac{1}{2}(n-3)(n-4),$$

according as the n points are required to be collinear, co-planar but not collinear, or not co-planar.

Accordingly, the necessary and sufficient condition (1) for a central configuration of the m_i , when expressed in terms of the $\frac{1}{2}n(n-1)$ distances $\rho_{ik} = \rho_{ki}$, must be written in all three cases (6₁)–(6₃) in the form

$$(7) \quad (JU^2)_{\rho_{ik}} + \sum_{s=1}^p \chi_s (R_s)_{\rho_{ik}} = 0, \quad 1 \leq i < k \leq n,$$

where $J, U; R_1, \dots, R_p$ are the functions (3₁)–(3₂); (5) of the ρ_{ik} , and χ_1, \dots, χ_p denote Lagrangian multipliers.

§358. As an application of this method, it will now be easy to determine all collinear central configurations belonging to $n = 3$ given positive numbers m_i .

In this case, the system (5) of p geometrical conditions reduces to a single equation $R = 0$, since (6₁) gives $p = 1$, if $n = 3$. Actually, this $R = 0$ is represented by $\rho_{13} = \rho_{12} + \rho_{23}$, if the notation is chosen so that m_2 lies between m_1 and m_3 on the line. Accordingly, $p = 1$ and $R = \rho_{13} - \rho_{12} - \rho_{23}$. Hence, if (i, j, k) denotes one of the three cyclic permutations of $(1, 2, 3)$, and χ the single Lagrangian multiplier $\chi_1 = \chi_p$, the necessary and sufficient condition (7) for a central configuration reduces to the three equations $(JU^2)_{\rho_{ik}} + \chi R_{\rho_{ik}} = 0$, where $R_{\rho_{ik}} = (-1)^j$, while J, U are given by (3₁)–(3₂); so that

$$(8) \quad \rho_{ik} U \mu^{-1} - \rho_{ik}^{-2} J + (-1)^j m_j K = 0, \quad \text{where } K = \chi: (2m_1 m_2 m_3 U).$$

Hence, $\det(\rho_{ik}, -\rho_{ik}^{-2}, (-1)^j m_j) = 0$, i.e.,

$$(9) \quad \begin{vmatrix} \rho_{23} & \rho_{23}^{-2} & + m_1 \\ \rho_{31} & \rho_{31}^{-2} & - m_2 \\ \rho_{12} & \rho_{12}^{-2} & + m_3 \end{vmatrix} = 0, \quad \text{where } \rho_{13} = \rho_{12} + \rho_{23}; \quad (\rho_{ik} = \rho_{ki}).$$

Conversely, if two given positive numbers ρ_{12}, ρ_{23} are such that (9) is satisfied, the computation of the minors of (9) shows that the three homogeneous linear equations (8) determine $U\mu^{-1}, -J, K$ up to a common factor in such a way that the value of the ratio $U\mu^{-1}:J$ is precisely that assigned by (3₁)–(3₃). Accordingly, the necessary and sufficient condition (7) is reduced to the determination of all pairs of positive numbers ρ_{12}, ρ_{23} which satisfy (9). Actually, the last remark of §355 shows that only the ratio of the two unknowns can be determined. Correspondingly, if one puts

$$(10) \quad \rho_{12}:\rho_{23} = \lambda(>0), \text{ then } \rho_{13}:\rho_{23} = 1 + \lambda, \text{ by } \rho_{13} = \rho_{12} + \rho_{23};$$

and so (9) is a condition for $\lambda = \lambda(m_1, m_2, m_3)$ alone. In fact, on multiplying the first and second columns of (9) by ρ_{23}^{-1} and ρ_{23}^{-2} , respectively, then using (10), and finally developing the resulting determinant, one sees that the condition (9) appears in the form

$$(11) \quad (m_2 + m_3)\lambda^5 + (2m_2 + 3m_3)\lambda^4 + (m_2 + 3m_3)\lambda^3 \\ - (3m_1 + m_2)\lambda^2 - (3m_1 + 2m_2)\lambda - (m_1 + m_2) = 0.$$

Consequently, the problem is reduced to the determination of the positive roots $\lambda = \lambda(m_1, m_2, m_3)$, if any, of the quintic equation (11).

It is easy to see that (11) has, for three arbitrarily given masses m_i , exactly one positive root λ , which is, in addition, such that

$$(12) \quad 0 < \lambda(m_1, m_2, m_3) \equiv \lambda \leq 1 \quad \text{according as} \quad m_1 \leq m_3.$$

In fact, since $(m_2 + m_3) > 0$, the quintic polynomial is positive for large positive λ ; while it attains at $\lambda = 0$ the negative value $-(m_1 + m_2)$. Since its value at $\lambda = 1$ is seen to be $7(m_3 - m_1) \geq 0$, it follows that there exists at least one root satisfying (12). On the other hand, there cannot exist more than one positive root, since the coefficients of (11) have only one change of sign.*

The dissymmetry of the three indices in (11), (12) is due, of course, to the corresponding dissymmetry in (10). In further agreement with (10) and (12), the relation (11) remains unchanged if one interchanges λ and $1/\lambda$ and, at the same time, m_1 and m_3 . The latter interchange is admissible, since the only normalization was that m_2 is

* Incidentally, the unique positive root is the only real root. In fact, (11) may be written in the form

$$(11 \text{ bis}) \quad \{-3\lambda^2 - 3\lambda - 1\}m_1 + \{(\lambda^3 - 1)(\lambda + 1)^2\}m_2 + \{\lambda^3(\lambda^2 + 3\lambda + 3)\}m_3 = 0$$

and the three coefficients $\{ \}$ of the $m_i > 0$ are seen to be negative for every $\lambda < 0$; so that no root $\lambda \leq 0$ is possible.

placed between m_1 and m_3 on the line. Since any of the three given m_i may be placed between the other two, and since each of these placements was shown to lead to exactly one determination of $\beta\rho_{12}$, $\beta\rho_{23}$, $\beta\rho_{13}$, where $\beta > 0$ is an arbitrary factor of proportionality (which, by the end of §395, may be omitted), the result of the above discussion may be summarized as follows:

To $n = 3$ arbitrarily given distinct masses m_i , there exist exactly three distinct collinear central configurations; furthermore, only two or all three of these central configurations are identical in the sense of §355 according as only two or all three of the values m_i of the given masses are equal.

§359. As another application of the criterion (7), it is easy to show that, for n arbitrarily given masses m_i which may or may not be distinct, the regular and only the regular tetrahedron is a non-flat central configuration for $n = 4$; that the equilateral and only the equilateral triangle is a non-collinear central configuration for $n = 3$; finally, that the segment is a central configuration for $n = 2$.

The assumptions of these three statements are the same, namely, that the non-negative integer (6_{n-1}) vanishes in the three respective cases $n = 4, 3, 2$; i.e., that the number of the geometrical conditions (5) is $p = 0$. Thus, (7) reduces to $(JU^2)_{\rho_{ik}} = 0$ and may, therefore, be written, in view of (3₁)–(3₂), as $\rho_{ik}^3 = \mu J/U$, where $(i, k) = (1, 2), \dots, (n-1, n)$. But these $\frac{1}{2}(n-1)n$ conditions, where $n = 4, 3, 2$, can be satisfied only if all $\frac{1}{2}(n-1)n [= 6, 3, 1]$ distances ρ_{ik} are equal; in which case (3₁) shows that $\rho_{ik}^3 = \mu J/U$ is actually satisfied, whether the given values of the n masses are distinct or not. Thus, the proof is complete.

§360. Apparently, these three cases, which are characterized by $p = 0$, exhaust all configurations which are central for arbitrary values of m_1, \dots, m_n , where n is unrestricted. Furthermore, the number $q = q(n; m_1, \dots, m_n)$ of all central configurations belonging to n given m_i is likely to be less than a bound q_n which is independent of the m_i ; while q_n itself remains bounded as $n \rightarrow \infty$.

The largest contribution to $q(n; m_1, \dots, m_n)$ seems to be due to the collinear central configurations. Actually, an enumeration of all $q(n; m_1, \dots, m_n)$ central configurations for arbitrary $n; m_1, \dots, m_n$ represents a fascinating unsolved problem which depends on a complete discussion of certain real algebraic equations.

(i) First, consider the problem of collinear central configurations

of n given m_i . If $n = 3$, the corresponding configurations are those enumerated in §358 and depend, therefore, on the given numbers m_1, m_2, m_3 , since so does the positive root λ of (11). In order to extend the method of §358 to any n , one can proceed by first choosing the notations so that m_j lies, for $j = 2, \dots, n-1$, between m_{j-1} and m_{j+1} on the line, and then expressing every ρ_{ik} , where $1 \leq i < k \leq n$, as the sum of $k-i$ of the successive distances $\rho_{l, l+1}$, where $l = 1, \dots, n-1$. This supplies between the ρ_{ik} the p [cf. (6₁)] geometrical conditions (5) which are, therefore, linear equations in the present case. However, application of the criterion (7) leads to a simultaneous system of $n-2$ non-linear algebraic equations, a system represented by (11) if $n = 3$. And what remains to be done is a discussion of this system with respect to reality and to its compatibility with the n given values $m_i > 0$; this problem of compatibility being represented, if $n = 3$, by the fact that the unknown ratio $\lambda = \rho_{12}:\rho_{23}$ must be positive. This discussion is a highly involved algebraic task, in which the difficulties of an explicit procedure increase rapidly with n . The literature of the subject contains a statement to the effect that to every numeration of n given masses (which may or may not be distinct) there exists on the line exactly one central configuration; so that, in particular, the number of distinct collinear central configurations belonging to n arbitrarily given distinct masses is $\frac{1}{2} \cdot (n)!$.

(ii) Next, consider the non-collinear flat case (6₂). If $n = 3$, the problem is solved completely by §359. If $n > 3$, the system (5) is no longer linear, since it is represented by (4) even in the lowest case, where $n = 4$, $p = 1$. In this case, application of the criterion (7) shows that the four sides and two diagonals of the quadrangle must satisfy, besides the geometrical identity (4), the necessary condition

$$(13) \quad (\rho_{12}^3 - \rho_{32}^3)(\rho_{13}^3 - \rho_{34}^3)(\rho_{14}^3 - \rho_{24}^3) = (\rho_{12}^3 - \rho_{24}^3)(\rho_{13}^3 - \rho_{32}^3)(\rho_{14}^3 - \rho_{43}^3),$$

which is, however, for four given m_i , not sufficient for the six conditions represented by (7), (4); in fact, the question of reality and compatibility still remains to be discussed. For instance, it is easily verified directly from (1) that the square is a central configuration only in case the four m_i are equal. And a detailed discussion shows that there belongs to $n = 4$ given m_i at least one non-collinear flat central configuration only when the m_i satisfy certain inequalities. The restrictions become, of course, stronger as n increases.

(iii) Finally, the non-flat central configurations are the scarcest. For not only is it likely that, except for the case $n = 4$, $p = 0$ treated completely in §359, one must then subject the n given m_i to at least one condition $f(m_1, \dots, m_n) = 0$; but, apparently, there exist only a finite number of integers n for which one can choose at least one set m_1, \dots, m_n possessing at least one non-flat central configuration. A corresponding conjecture cannot be correct in the non-collinear flat case, since n equal m_i , when placed at the corners of a regular n -gon, clearly form a central configuration for any n . One can also place $n - 1$ equal m_i at the corners of a regular $(n - 1)$ -gon and an arbitrary n -th mass at the mid-point. Furthermore, one can combine, under obvious restrictions, either of these flat models with its polar model. Now, there clearly exist corresponding models in the non-flat case of regular polyhedra, in which case, however, these constructions are possible only for a finite set of particular values of n .

§361. The notion of a central configuration will now be applied to a surprising analysis of the ultimate shrinking process of the configuration formed by n arbitrarily given bodies m_i at a given date t , when the solution under consideration leads to a simultaneous collision of all n bodies as t tends to some t^0 (cf. §335).

It will be shown that if t is very close to t^0 , the configuration belonging to t is very close to a central configuration of the n moving m_i , where it is understood that only the relative magnitudes and the relative locations of the shrinking mutual position vectors $\xi_i - \xi_k$ of the n moving m_i are to be considered (cf. the end of §355). This asymptotic description of any possible simultaneous collision is, perhaps, the deepest among all the known local theorems in the problem of n bodies.

The proof will require the Tauberian refinement (18₁), §337 of (16₂), §335 bis, as well as another, more primitive, Tauberian fact. The latter may be described as follows:

§362. Denote by dots differentiations with respect to an independent variable u , where $0 < u < +\infty$, and let $g(u)$ be a function for which there exists a continuous $\dot{g}(u)$ and a finite limit $g(+\infty)$. Then, though $g(u)$ is, for $u \rightarrow +\infty$, asymptotically equal to the constant $g(+\infty)$, obvious examples show that differentiation of this asymptotic relation is not, in general, admissible, i.e., that $\dot{g}(u)$ need not tend to 0 as $u \rightarrow +\infty$. But if there exists a continuous $\ddot{g}(u)$, the boundedness of this second derivative is a Tauberian condition

in the sense of §336. In other words, if $g(u)$ tends to a finite limit and $|\dot{g}(u)| < \text{const.}$ as $u \rightarrow +\infty$, then $\dot{g}(u) \rightarrow 0$.*

§363. It can be assumed without loss of generality that the given simultaneous collision of the n bodies takes place at $t^0 = 0$, while t tends decreasingly to $t^0 = 0$. Then $t - t^0 = t > 0$. It is easily verified that the asymptotic relations (17₁), (17₂), (18₂) of §336-§337 may respectively be written as

$$(14_1) \quad t^{-\frac{2}{3}}J \rightarrow \mathfrak{u}_0 > 0; \quad (14_2) \quad t(t^{-\frac{2}{3}}J)' \rightarrow 0; \quad (14_3) \quad t^2(t^{-\frac{2}{3}}J)'' \rightarrow 0,$$

where \mathfrak{u}_0 denotes the constant $(\frac{3}{2}\mu_0^{\frac{1}{3}})^{\frac{2}{3}}$, and $t \rightarrow +0$. Since $J = \sum m_i \xi_i^2$, it is indicated by (14₁) that, as $t \rightarrow +0$, the n bodies collide at the origin $\xi = 0$ of the barycentric coordinate system ξ in such a way that the linear dimensions of the configuration formed by the n bodies for a small t are nearly proportional to $t^{\frac{1}{3}}$. Thus, it will be convenient to magnify the unit of length in the proportion $1:t^{\frac{1}{3}}$, by considering

$$(15_1) \quad \xi_i = t^{-\frac{1}{3}}\xi_i, \quad \varrho_{ik} = t^{-\frac{1}{3}}\rho_{ik}; \quad (15_2) \quad \mathbf{J} = t^{-\frac{2}{3}}J, \quad \mathbf{U} = t^{\frac{2}{3}}U$$

instead of ξ_i , $\rho_{ik} = |\xi_i - \xi_k|$, $J = \sum m_i \xi_i^2$, $U = \sum^* m_j m_k / \rho_{jk}$, respectively. Then the exact formulation of the statement of §361 is that, while (1) need not hold for a fixed $t \neq 0$, one has, as $t \rightarrow +0$,

$$(16) \quad (\mathbf{J}\mathbf{U}^2)_{\xi_i} \rightarrow 0; \quad i = 1, \dots, n.$$

In fact, the last remark of §355 implies that the change of scale introduced by the substitutions (15₁)-(15₂) is immaterial for what has to be proved.

§364. First, it will be shown that, as $t \rightarrow +0$,

$$(17_1) \quad \frac{2}{3}\mathbf{J} - \mathbf{U} \rightarrow 0; \quad (17_2) \quad \varrho_{ik} > \text{const.} > 0.$$

* In the proof of this Tauberian lemma, it may obviously be assumed that g is a scalar. Then there cannot exist a pair of sufficiently small positive numbers η , δ such that $|\dot{g}(u)| > \eta$ holds at every point of infinitely many u -intervals which have the common length δ and cluster at $u = +\infty$. For otherwise $g(u)$ would vary on each of these intervals by an amount $\geq \eta \cdot \delta$; and so $g(u)$ could not tend to a finite limit as $u \rightarrow +\infty$. Consequently, there exists for every $\epsilon > 0$ an $N = N_\epsilon$ such that any u -interval which has the length ϵ and is contained in the region $N_\epsilon \leq u < +\infty$ contains at least one point u at which $|\dot{g}(u)| < \epsilon$. Since the assumption $|\ddot{g}(u)| < \text{const.}$ implies that $\dot{g}(u)$ cannot vary on intervals of length ϵ by more than $\epsilon \cdot \text{const.}$, it follows that $|\dot{g}(u)| < \epsilon + \epsilon \cdot \text{const.}$ whenever $u > N_\epsilon$. This proves that $\dot{g}(u) \rightarrow 0$ as $u \rightarrow +\infty$.

To this end, it will be convenient to replace t by $\mathfrak{t} = -\log t$; so that $\mathfrak{t} \rightarrow +\infty$ as $t \rightarrow 0$, and

$$(18_1) \quad t = \exp(-\mathfrak{t}); \quad (18_2) \quad tf' = -\dot{f}, \quad t^2 f'' = \ddot{f} + \dot{f},$$

where the primes and dots denote differentiations with respect to t and \mathfrak{t} , respectively. For instance, it is easily verified from (18₁)–(18₂) and (15₁) that the equation of motion, $m_i \xi_i'' = U_{\xi_i}$, may be written as

$$(19_1) \quad m_i(\ddot{\xi}_i - \tfrac{1}{3}\dot{\xi}_i - \tfrac{2}{9}\xi_i) = U_{\xi_i};$$

$$(19_2) \quad \mathbf{U} = \sum^* m_j m_k / \varrho_{jk}, \quad \varrho_{jk} = |\xi_j - \xi_k|,$$

since $\mathbf{U} = t^3 U$, $U_{\xi_i} = t^3 U_{\xi_i}$. Similarly, it is seen from (15₁)–(15₂), (18₁)–(18₂) that the energy integral $\frac{1}{2} \sum m_i \dot{\xi}_i'^2 - U = h$ and its equivalent formulation $J'' = 2U + 4h$ may be written as

$$(20_1) \quad \tfrac{1}{2} \sum m_i (\dot{\xi}_i - \tfrac{2}{3}\xi_i)^2 - \mathbf{U} = h \exp(-\tfrac{2}{3}\mathfrak{t});$$

$$(20_2) \quad \ddot{\mathbf{J}} - \tfrac{5}{3}\dot{\mathbf{J}} + \tfrac{4}{9}\mathbf{J} = 2\mathbf{U} + 4h \exp(-\tfrac{2}{3}\mathfrak{t}).$$

Finally, application of (18₂) to the function $f = \mathbf{J}$ defined by (15₂) shows that (14₁), (14₂), (14₃) are equivalent to

$$(21_1) \quad \mathbf{J} \rightarrow \mathfrak{u}_0 > 0; \quad (21_2) \quad \dot{\mathbf{J}} \rightarrow 0; \quad (21_3) \quad \ddot{\mathbf{J}} \rightarrow 0,$$

where the arrows refer to $t \rightarrow t^0 + 0 = +0$, i.e., to $\mathfrak{t} \rightarrow +\infty$.

On letting $\mathfrak{t} \rightarrow +\infty$ in (20₂), where $h = \text{const.}$, one sees that (17₁) is implied by (21₁)–(21₃). On the other hand, (17₁) and (21₁) show that \mathbf{U} tends to a finite limit; so that (17₂) follows from (19₂).

Next, it will be shown that, as $t \rightarrow +0$, i.e., as $\mathfrak{t} \rightarrow +\infty$,

$$(22_1) \quad \xi_i \rightarrow 0; \quad (22_2) \quad |\dot{\xi}_i| < \text{Const.}; \quad (22_3) \quad |\ddot{\xi}_i| < \text{Const.}$$

To this end, notice first that from (15₁)–(15₂), where $J = \sum m_i \xi_i'^2$, one has $\mathbf{J} = \sum m_i \xi_i'^2$; hence, $\dot{\mathbf{J}} = 2 \sum m_i \xi_i \dot{\xi}_i$. Consequently, on letting $\mathfrak{t} \rightarrow +\infty$ in (20₁), one sees from (21₂) and (17₁) that $\sum m_i \dot{\xi}_i'^2 \rightarrow 0$. This proves (22₁). Furthermore,

$$(23_1) \quad |\xi_i| < \text{const.}; \quad (23_2) \quad |\mathbf{U}_{\xi_i}| < \text{Const.}$$

In fact, (23₁) is clear from (21₁), since $\mathbf{J} = \sum m_i \xi_i'^2$; while (23₂) is implied by (19₂) and (17₂). Now, (22₂) is clear from (23₁)–(23₂), (22₁) and (19₁). Finally, on differentiating (19₁) with respect to \mathfrak{t} and then using (22₁)–(22₂), one sees that in order to prove (22₃), it is sufficient to show that the partial derivatives of the second order of the function $\mathbf{U}(\xi_1, \dots, \xi_n)$ remain bounded as $\mathfrak{t} \rightarrow +\infty$. But

(23₁), (19₂), (17₂) clearly imply the boundedness of these derivatives also.

According to (22₁) and (22₃), the Tauberian lemma of §362 is applicable to $g(u) = \xi_i$, where $u = t$. Hence, not only $\xi_i \rightarrow 0$ but also $\dot{\xi}_i \rightarrow 0$. It follows, therefore, from (19₁) that $\frac{2}{3}m_i\xi_i + U_{\xi_i} \rightarrow 0$. Since $\mathbf{J} = \sum m_i \xi_i^2$, this may be written as $\frac{1}{3}\mathbf{J}_{\xi_i} + U_{\xi_i} \rightarrow 0$. This relation, when combined with (21₁) and (17₁), completes the proof of (16).

§365. The interpretation of (16) in terms of (1), as given in §361, was very cautious. For all that was said is that if t is very close to the date t^0 of simultaneous collisions, the configuration is very close to a central configuration of the given m_i . This does not imply that, as $t \rightarrow t^0$, the configuration must tend to a central configuration of the given m_i . For, as far as present knowledge goes, it would be possible that the configuration comes closer and closer to more than one central configuration of the m_i in such a way as to oscillate between these central configurations, as $t \rightarrow t^0$. Of course, this possibility cannot occur unless the n given m_i determine infinitely many central configurations which are distinct in the sense defined at the end of §355. In §360, it appeared to be a reasonable conjecture that such is never the case, i.e., that the integer $q(n; m_1, \dots, m_n)$ defined at the beginning of §360 always exists. But no proof is known for the truth of this hypothesis.

§365 bis. For those n and m_i for which $q(n; m_1, \dots, m_n) < +\infty$ is established, it follows, of course, that the configuration must tend to a well-determined central configuration of the m_i , as $t \rightarrow t^0$. Hence, if $q(n; m_1, \dots, m_n) < +\infty$ is established for the given m_i , then—and only in this case—one can infer from (21₁) the existence of the $\frac{1}{2}(n-1)n$ limits

$$(24) \quad 0 < {}^0\varrho_{ik} = \lim \varrho_{ik} < +\infty, \quad \text{where} \quad \varrho_{ik} = t^{-\frac{1}{2}}\rho_{ik}; \quad t \rightarrow 0.$$

In fact, $\mathbf{J} \equiv t^{-\frac{1}{2}}J$ may be written, by §322 bis, as $\mathbf{J} = \sum^* m_j m_k \varrho_{jk}^2 / \sum m_i$ where the $\varrho_{jk} = |\xi_j - \xi_k|$ remain bounded, by (23₁), and cannot tend to 0, by (17₂). Incidentally, $\mathbf{u}_0 = \sum^* m_j m_k {}^0\varrho_{jk}^2 / \sum m_i$, by (21₁), while $\frac{2}{3}\mathbf{u}_0 = \sum^* m_j m_k / {}^0\varrho_{jk}$, by (17₁) and (19₂); so that

$$(25_1) \quad \frac{4}{27}\mathbf{u}_0^3 \sum m_i = (\sum^* m_j m_k {}^0\varrho_{jk}^2) (\sum^* m_j m_k / {}^0\varrho_{jk})^2;$$

$$(25_2) \quad \mathbf{u}_0 \sum m_i = \sum^* m_j m_k {}^0\varrho_{jk}^2.$$

§366. Notice that the explicit representation (25₁) of \mathbf{u}_0 contains only the m_i and the ratios ${}^0\varrho_{jk} : {}^0\varrho_{rs}$ of the $\frac{1}{2}(n-1)n$ limits (24); ratios which are algebraic functions of the m_i , since the limits (24) are mutual distances in a central configuration belonging to the m_i . Thus, while the last remark of §355 leaves the limits (24) undetermined with respect to a common positive factor, it is seen from (25₁) and (25₂) that this factor of proportionality is, in the present case, uniquely determined by the m_i and the central configuration, since (25₁) determines \mathbf{u}_0 as a function of the m_i .

Since \mathbf{u}_0 was introduced into (14₁) as $(\frac{3}{2}\mu_0^{\frac{1}{2}})^{\frac{4}{3}}$, there also follows, in terms of the m_i , a determination of the positive constant μ_0 whose existence was established in §335 bis.

§367. As an illustration, consider the case of $n = 3$ arbitrary masses m_i . In this case, the assumption of §365 bis is satisfied, since $q(3; m_1, m_2, m_3) \leq 4$ for arbitrary m_i . In fact, there exists, by §359, only one non-collinear central configuration, namely, the equilateral triangle; while, by the end of §358, the number of collinear central configurations is equal to the number of the distinct m_i . Thus, if $n = 3$, the difficulty pointed out in §365 does not arise, and so §365 bis is applicable.

§368. It is natural to ask whether or not every simultaneous collision of n bodies must take place in such a way that the n barycentric initial position vectors $\xi_i(t)$ tend to their common limit 0 in definite directions, that is to say so that all of the vectors $\xi_i(t)/|\xi_i(t)|$ of unit length have limits. It was shown in §351 that in case of a binary collision the answer to the corresponding question is affirmative. In case of a simultaneous collision of all $n (> 2)$ bodies, it seems to be much more difficult to prove that the bodies cannot move, before colliding at the centre of mass, in spirals without asymptotes.

§368 bis. On the other hand, it is easy to see that if the simultaneous collision is such that there exist limiting positions for the tangents of the n paths $\xi_i = \xi_i(t)$, then, whether the condition $q(n; m_1, \dots, m_n) < +\infty$ of §365 is satisfied or not, the configuration must tend to a definite central configuration, in the sense that all $\frac{1}{2}n(n-1)$ limits (24) exist. For in the case of definite limiting directions, one easily infers from (15₁), (18₁)–(18₂) and (22₁) that there exist finite limits $\lim t^{\frac{1}{3}}\xi'_i = \frac{2}{3}\lim \xi_i$, where at least $n-1$ of the n limits $\lim \xi_i$ do not vanish, since the origin is centre of mass. But $\varrho_{ik} = |\xi_i - \xi_k|$, so that also the limits (24) exist.

Homographic Solutions

§369. A given solution $\xi_i = \xi_i(t)$ of the problem of n bodies is called homographic if the configuration formed by the n bodies at a given t moves in the inertial barycentric coordinate system ξ in such a way as to remain similar to itself when t varies. By this is meant that there exist a scalar $r = r(t) \geq 0$, an orthogonal 3-matrix $\Omega = \Omega(t)$ and a 3-vector $\tau = \tau(t)$ such that for every i and t one has $\xi_i = r\Omega\xi_i^0 + \tau$, where ξ_i , r , Ω , τ belong to an arbitrary t and ξ_i^0 denotes ξ_i at some initial $t = t^0$. Actually, only the dilatation and rotation, represented by the unknowns $r = r(t)$ and $\Omega = \Omega(t)$, are possible, since the translation vector $\tau = \tau(t)$ must vanish identically in view of the barycentric condition $\sum m_i \xi_i = 0$.

Needless to say, the homographic solutions are of a rather restricted type, since the system $m_i \xi_i'' = U_{\xi_i}$ of order $6n$ has to be satisfied by the $1 + 3$ scalar functions and the $3n$ integration constants which are represented by $r(t)$, $\Omega(t)$ and the ξ_i^0 , respectively.

§370. First, a few identities will be collected.

According to §369, an homographic solution $\xi_i(t)$ is characterized by the existence of a rotation $\Omega(t)$ and a dilatation $r(t) > 0$ such that, for every i and t ,

$$(1) \quad \xi_i = r\Omega\xi_i^0, \quad \text{i.e.} \quad x_i = r\xi_i^0, \quad \text{where} \quad x = \Omega^{-1}\xi$$

is a barycentric, but not necessarily inertial, coordinate system, and the superscripts 0 always refer to a fixed initial date t^0 .

For instance, it is clear from (1) that

$$(2_1) \quad r^0 \equiv r(t^0) = 1, \quad (r = r(t) > 0); \quad (2_2) \quad \Omega^0 \equiv \Omega(t^0) = E,$$

where E denotes the unit 3-matrix. It is also seen from (1) that

$$(3_1) \quad J = J^0 r^2; \quad (3_2) \quad U = U^0 / r; \quad (3_3) \quad \Omega^{-1} U_{\xi_i} = U_{\xi_i}^0 / r^2,$$

since the scalars $J = \sum m_i \xi_i^2$, $U = \sum^* m_j m_k / |\xi_j - \xi_k|$ are for every t invariant, hence their gradients covariant, under the rotation Ω . Similarly, Σ^0 will denote the matrix formed by the initial values of the three scalars $s_\nu = s_\nu(t)$ which are defined in terms of $\Omega = \Omega(t)$ by (5), §66; so that

$$(4_1) \quad \Sigma = \begin{pmatrix} 0 & -s_3 & s_2 \\ s_3 & 0 & -s_1 \\ -s_2 & s_1 & 0 \end{pmatrix};$$

$$(4_2) \quad \Sigma^2 = \begin{pmatrix} -s_2^2 - s_3^2 & s_1 s_2 & s_1 s_3 \\ s_2 s_1 & -s_3^2 - s_1^2 & s_2 s_3 \\ s_3 s_1 & s_3 s_2 & -s_1^2 - s_2^2 \end{pmatrix}$$

(cf. (5)–(6), §66). Since from (1), where $\xi_i^0 = \text{const.}$, one has

$$x_i = r \xi_i^0, \quad x_i' = r' \xi_i^0 \equiv r' E \xi_i^0, \quad x_i'' = r'' \xi_i^0 \equiv r'' E \xi_i^0,$$

and since the definition $x = \Omega^{-1} \xi$ of the rotating coordinate system x may be written in the form (8), §69 by placing $\Xi = \xi$, $X = x$, one sees from (10₁)–(10₂), §69 that

$$(5_1) \quad \Omega^{-1} \xi_i' = (r' E + r \Sigma) \xi_i^0;$$

$$(5_2) \quad \Omega^{-1} \xi_i'' = \{r'' E + 2r' \Sigma + r(\Sigma' + \Sigma^2)\} \xi_i^0.$$

It is clear from (5₂) and (3₃) that if $m_i a_i$ denotes the constant 3-vector $U_{\xi_i}^0$, then along the homographic solution $\xi_i = \xi_i(t)$ of $m_i \xi_i'' = U_{\xi_i}$ one has

$$(6_1) \quad K(t) \xi_i^0 = a_i; \quad (6_2) \quad r^2 \{r'' E + 2r' \Sigma + r(\Sigma' + \Sigma^2)\} \equiv K = (\kappa_{pq}),$$

(6₂) being the definition of a 3-matrix function $K = (\kappa_{pq})$ of t . If A' denotes, as in §1, the transposed of a matrix A , then obviously $E = E'$, while (4₂), (4₁) show that $(\Sigma^2)' = \Sigma^2$, $\Sigma' = -\Sigma$; $(\Sigma')' = -\Sigma'$; hence, from (6₂),

$$(7_1) \quad \frac{1}{2}(K + K') = r^2(r'' E + r \Sigma^2); \quad (7_2) \quad \frac{1}{2}(K - K') = r^2(r \Sigma' + 2r' \Sigma).$$

The above formulae allow an essential simplification in the special case in which the particular solution $\xi_i = \xi_i(t)$ is planar in the sense of §324. For then the barycentric inertial coordinate system ξ may be so chosen that the third component of each of the 3-vectors $\xi_i(t)$ vanishes identically, i.e., that Ω is given by (13₁), §72, where $\phi' = \phi'(t)$ denotes the angular velocity of the rotating coordinate system $x = \Omega^{-1} \xi$. Hence, it is easily verified from (5₁) that $\xi_i'^2 = (r' \xi_i^0)^2 + (r \phi' \xi_i^0)^2$, and that the components of the 3-vector $\xi_i \times \xi_i'$ are, in view of (1), equal to 0, 0, $\phi'(r \xi_i^0)^2$, respectively. It follows, therefore, from (3₁), where $J = \sum m_i \xi_i^2$, that the kinetic energy $T = \frac{1}{2} \sum m_i \xi_i'^2$ and the 3-vector integral $\sum m_i \xi_i \times \xi_i' = C$ reduce to

$$(8_1) \quad T = \frac{1}{2}(r'^2 + r^2 \phi'^2) J^0; \quad (8_2) \quad \phi' r^2 J^0 = |C|, \quad (J^0 > 0),$$

if one chooses the alternative sign in (6), §323 so that $\phi' \geq 0$. Finally, since $s_1 \equiv 0$, $s_2 \equiv 0$, $s_3 = \phi'$ by §72, substitution of (4₁)–(4₂) into (6₂) shows that

$$(9) \quad K(t) \equiv K = \begin{pmatrix} r^2(r'' - r\phi'^2) & -r^2(r\phi'' + 2r'\phi') & 0 \\ r^2(r\phi'' + 2r'\phi') & r^2(r'' - r\phi'^2) & 0 \\ 0 & 0 & r^2r'' \end{pmatrix}.$$

In what follows, non-planar homographic solutions will not be excluded; so that only (1)–(7₂), but not (8₁)–(9), are applicable. Also in this general case, the equivalent formulation $J'' = 2U + 4h$ of the energy integral $T - U = h$ may be written, by (3₁)–(3₂), as

$$(10) \quad (rr'' + r'^2)J^0 - r^{-1}U^0 = 2h, \quad (J^0 > 0, U^0 > 0).$$

§370 bis. There are two limiting types of homographic solutions.

On the one hand, it is possible that the configuration is dilating without rotation, i.e., that $\Omega(t) \equiv E$. These particular homographic solutions are, in view of (1), characterized by

$$(11) \quad \xi_i = r\xi_i^0, \quad \text{i.e.,} \quad x_i \equiv \xi_i, \quad (\Omega(t) \equiv E, r = r(t) > 0),$$

and will be called homothetic solutions.

On the other hand, it is possible that the configuration is rotating without dilatation, i.e., that $r(t) \equiv 1$. These particular homographic solutions are, in view of (1), characterized by

$$(12) \quad \xi_i = \Omega\xi_i^0 \quad \text{i.e.,} \quad x_i \equiv \xi_i^0, \quad (r(t) \equiv 1, \Omega = \Omega(t)),$$

and will be called solutions of relative equilibrium. This name is justified by the fact that, in the case (12) and only in this case, each of the n particles appears at rest in a rotating barycentric coordinate system x (i.e., $x_i(t) \equiv \text{const.}$). This is possible only when the forces of gravitation acting between the m_i are at every t in exact balance with the apparent forces (§318 bis) introduced by the rotation of the system x .

It is clear that a solution of relative equilibrium cannot be homothetic. A general homographic solution satisfies neither (11) nor (12). A description of these two particular types may be given by the following facts, which will be proved in §372.

(I) An homographic solution is homothetic if and only if it has no invariable plane (i.e., if and only if $C = 0$).

(II) An homographic solution is a solution of relative equilibrium if and only if it is planar and rotates with a constant angular velocity ($\neq 0$).

According to (I) and §329–§331, every collinear but not rectilinear solution is homographic but not homothetic, while the homothetic collinear solutions are identical with those rectilinear solutions which are homographic. It is also clear that in order that a collinear solution be a solution of relative equilibrium, it is necessary (but, by (II), not sufficient) that the solution be not rectilinear.

§371. In the terminology of §324–§325, there will be proved in §373–§374 the following facts, which lie deeper than (I)–(II), §370 bis:

- (i) If an homographic solution is not flat, then it is homothetic.
- (ii) If an homographic solution is flat, then it is planar.

The converse of (i) is not true, since there exist planar (and even rectilinear) homothetic solutions. This, when combined with (ii), may be expressed by saying that every homographic solution is either planar or homothetic but may be both.

If an homographic solution is not planar, then (i)–(ii) assure that (11) is valid; so that the kinetic energy $T = \frac{1}{2} \sum m_i \dot{\xi}_i^2$ reduces to

$$\frac{1}{2} \sum m_i (r' \xi_i^0)^2 \equiv \frac{1}{2} r'^2 J^0.$$

Since (8₁) holds in the planar case and (3₂) in every case, it follows that the energy integral $T - U = h$ of every homographic solution may be written in the form

$$(13) \quad \frac{1}{2} (r'^2 + r^2 \phi'^2) J^0 - r^{-1} U^0 = h,$$

if $\phi' = \phi'(t)$, which is defined as the angular velocity of the rotating coordinate system $x = \Omega^{-1} \xi$ in the planar case, is defined by $\phi'(t) \equiv 0$ in the non-planar case. In this sense, (8₂) holds in the non-planar case also, since then $C = 0$, by (I) and (i)–(ii). Finally, one sees from (6₂) that (9) holds with $\phi' \equiv 0$ if $\Sigma = 0$ for every t , which means, by the end of §69, that $\Omega(t) = \text{const.}$ Since (i)–(ii) show that $\Omega(t) = \text{const.}$ is satisfied in the non-planar case, it follows that (9) becomes valid for this case by placing again $\phi' \equiv 0$.

§372. The object of this article is to show that (I)–(II) are implied by (i)–(ii); while (i)–(ii) will be proved in the next two articles.

If an homothetic solution is planar, then (8₂) is applicable and shows that $C = 0$ if and only if the angular velocity $\phi'(t) \equiv 0$, and that $r(t) = \text{const.} (> 0)$ if and only if $\phi'(t) \equiv \text{const.} \neq 0 \neq |C|$.

This proves (I)–(II) for the planar case. If an homographic solution is not planar, then it is, by (i)–(ii), homothetic, and so, by §370 bis, certainly not a solution of relative equilibrium. This completes the proof of (II) and also shows that in order to complete the proof of (I), it is sufficient to prove that $C = 0$ for every non-planar homographic solution. Now, whether an homographic solution $\xi_i = \xi_i(t)$ is or is not planar, every term of the sum $C = \sum m_i \xi_i \times \xi_i'$ vanishes for every t , since $\zeta \times \zeta \equiv 0$, while $\xi_i = r\xi_i^0$, $\xi_i' = r'\xi_i^0$, by (11).

§373. The object of this article is to prove (i), §371.

Let $\xi_i = \xi_i(t)$ be a given non-flat homographic solution. Then not all n initial position vectors ξ_i^0 are co-planar, and so one can select three values of i , say $i = \alpha, \beta, \gamma$, such that $\det(\xi_\alpha^0, \xi_\beta^0, \xi_\gamma^0) \neq 0$. Thus, the 3-matrix $(\xi_\alpha^0, \xi_\beta^0, \xi_\gamma^0)$, which is independent of t , has a reciprocal matrix. Hence, application of (6₁) to $i = \alpha, \beta, \gamma$ shows that the 3-matrix $K(t)$ is the product of this reciprocal matrix and of the 3-matrix $(a_\alpha, a_\beta, a_\gamma)$, which, by the definition of the a_i (§370), again is independent of t ; so that, from (7₁)–(7₂),

$$(14_1) \quad r^2 r'' E + r^3 \Sigma^2 = \text{const.}; \quad (14_2) \quad r^3 \Sigma' + 2r^2 r' \Sigma = \text{Const.}$$

Since E is the unit matrix, $r^2 r'' E$ is a diagonal matrix in which all diagonal elements are equal. Hence, (14₁) implies that, on the one hand, those elements of the 3-matrix $r^3 \Sigma^2$ which are not diagonal elements and, on the other hand, the differences of any two of the diagonal elements of this 3-matrix $r^3 \Sigma^2$ are independent of t . This, when compared with (4₂), shows that both $r^3 s_\mu s_\nu$ and $r^3 (s_\mu^2 - s_\nu^2)$ are independent of t , where $(\mu, \nu) = (1, 2), (2, 3), (3, 1)$. Consequently, $r^3 s_\lambda^2$ is independent of t , where $\lambda = 1, 2, 3$. It follows, therefore, from (4₁) that Σ and r depend on t in such a way that $\Sigma = r^{-\frac{1}{3}} \Sigma_0$, where Σ_0 is a constant skew-symmetric matrix.

Hence, there exists a constant orthogonal matrix P_0 for which $P_0 \Sigma_0 P_0^{-1}$ becomes a skew-symmetric 3-matrix in which all elements of the third row vanish (cf. the beginning of §75). Since $\Sigma = r^{-\frac{1}{3}} \Sigma_0$, where $r = r(t)$ is a scalar and $\Sigma = \Sigma(t)$, it follows that all elements of the third row of the skew-symmetric matrix $P_0 \Sigma(t) P_0^{-1}$ vanish for every t . It follows, therefore, from §74 that the rotation $\Omega = \Omega(t)$ is one about an axis which has an invariable position with reference to the barycentric inertial coordinate system $\xi = (\xi^I, \xi^{II}, \xi^{III})$. Hence, §318 shows that this axis may be chosen to be the ξ^{III} -axis. Then the rotation $\Omega = \Omega(t)$ is given by (13₁), §72. Hence, (13₃), §72 shows that $s_1 \equiv 0, s_2 \equiv 0, s_3 = \phi'$, where $\phi' = \phi'(t)$ is the velocity of

rotation. Consequently, the proof of (i), §371 will be complete if one shows that $\phi'(t) \equiv 0$. For then there is no rotation at all, which means, by the definition of §370 bis, that the solution is homothetic.

In the proof of $\phi'(t) \equiv 0$, use will be made of the relation $r^3\phi'^2 = \text{const.}$, which is, in view of (13₃), §72, equivalent to the above result according to which all three $r^3s_\lambda^2$ are independent of t .

Since $\Omega = \Omega(t)$ is given by (13₁), §72, one can write (1), §370 as

$$\begin{aligned}\xi_i^{\text{I}} &= r(\xi_i^{0\text{I}} \cos \phi - \xi_i^{0\text{II}} \sin \phi), & \xi_i^{\text{II}} &= r(\xi_i^{0\text{I}} \sin \phi + \xi_i^{0\text{II}} \cos \phi), \\ \xi_i^{\text{III}} &= r\xi_i^{0\text{III}},\end{aligned}$$

where $(\xi_i^{\text{I}}, \xi_i^{\text{II}}, \xi_i^{\text{III}}) \equiv \xi_i(t)$ and $\xi_i^{0\nu} = \xi_i^\nu(t^0) = \text{const.}$ Hence, the third of the equations (5), §322 readily reduces to $C^{\text{III}} = r^2\phi'c$, where c denotes the constant $\sum m_i \{ (\xi_i^{0\text{I}})^2 + (\xi_i^{0\text{II}})^2 \}$. Thus, if $c = 0$, then all $\xi_i^{0\text{I}} = 0$ and all $\xi_i^{0\text{II}} = 0$, and so, by the above representation of the $\xi_i = \xi_i(t)$, all n bodies m_i are situated on the ξ^{III} -axis for every t . Since this contradicts the assumption that the given solution $\xi_i = \xi_i(t)$ is not flat, it follows that $c \neq 0$.

Hence, $C^{\text{III}} = r^2\phi'c$ may be divided by c and implies, therefore, that $r^2\phi' = \text{Const.}$ On comparing this with the relation $r^3\phi'^2 = \text{const.}$, found above, one sees that either $\phi' \equiv 0$ or the function $r = r(t)$, which is positive by (2₁), is independent of t . Thus, the proof of $\phi'(t) \equiv 0$ will be complete if one shows that the assumption $r = \text{const.}$ leads to a contradiction.

If $r = \text{const.}$, then (6₂) reduces to $K = r^3(\Sigma' + \Sigma^2)$, where $\det(\Sigma' + \Sigma^2) \equiv 0$, since, in view of $s_1 \equiv 0$, $s_2 \equiv 0$, all elements of the third column of either of the matrices (4₁)–(4₂) vanish identically; so that $\det K \equiv 0$. On the other hand, it was shown before (14₁)–(14₂) that K is the product of two constant matrices which have non-vanishing determinants; so that $\det K \neq 0$. Since this is a contradiction, the proof is complete.

§373 bis. One might have the impression that this complicated proof of (i), §371 is unnecessary, since the statement seems to be intuitive enough to be a direct consequence of the conservation of the angular momentum alone.

Such is, however, not the case. For otherwise (i), §371 would be true also in case the attraction is chosen to be inversely proportional to the third, instead of the second, power of the distance. But in this case, $r^2\{ \quad \}$ in (6₂) must be replaced by $r^3\{ \quad \}$, and so the relation $r^3\phi'^2 = \text{const.}$, found in §373, by $r^4\phi'^2 = \text{const.}$ And this is the

same condition as the relation $r^2\phi' = \text{Const.}$, found in §373 as a consequence of the conservation of angular momentum; so that this time there is only one relation between r and ϕ' , and so the proof breaks down.

Actually, the theorem itself is false. In other words, the problem of $n \geq 4$ bodies belonging to the inverse cubic law of gravitation possesses non-flat solutions which are homothetic but not homographic. An example to this effect may readily be obtained by adapting, to the case of $\frac{1}{2}n = 2$ congruent pairs of $n = 4$ masses, initial positions and initial velocities, the explicit calculations of the isosceles solutions ($n = 3$) which will be derived in §374 bis.

§374. The object of this article is to prove (ii), §371.

For the collinear case, (ii), §371 was already proved in §329. Let, therefore, $\xi_i = \xi_i(t)$ be a given homographic solution which is flat but not collinear. Then there exist among the n initial position vectors ξ_i^0 at least two, say ξ_α^0 and ξ_β^0 , such that $\xi_\alpha^0 \times \xi_\beta^0 \neq 0$. Since the solution is flat, all n initial vectors ξ_i^0 lie in one and the same plane through the origin of the inertial barycentric coordinate system $\xi = (\xi^I, \xi^{II}, \xi^{III})$. Hence, §318 shows that this plane may be chosen to be the (ξ^I, ξ^{II}) -plane. Then $\xi_i^{0III} = 0$ for every i . Hence, on denoting by $a_i^I, a_i^{II}, a_i^{III}$ the components of the constant 3-vector a_i , one can write (6₁), where $K = (\kappa_{pq})$, in the form

$$(15) \quad \kappa_{\nu 1}(t)\xi_i^{0I} + \kappa_{\nu 2}(t)\xi_i^{0II} = a_i^\nu, \quad (\nu = I, II, III),$$

where $i = 1, \dots, n$. Finally, all $a_i^{III} = 0$. In fact, a_i was introduced into (6₁) by the definition that $m_i a_i$ is the force of gravitation acting on m_i at the date $t = t^0$. And these forces cannot have, at the date $t = t^0$, components parallel to the ξ^{III} -axis, since all gravitating masses lie in the (ξ^I, ξ^{II}) -plane when $t = t^0$.

On applying (15) to $i = \alpha$ and $i = \beta$ and keeping $\nu (= I, II, III)$ fixed, one obtains for the two scalars $\kappa_{\nu 1}(t), \kappa_{\nu 2}(t)$ two linear equations which have constant coefficients and, since $\xi_\alpha^0 \times \xi_\beta^0 \neq 0$ and $\xi_i^{0III} = 0$, a non-vanishing determinant. Consequently, the two scalars $\kappa_{\nu 1}(t), \kappa_{\nu 2}(t)$ are linear combinations, with constant coefficients, of the two scalars $a_\alpha^\nu, a_\beta^\nu$, where α, β are the particular i -values selected above. Since the a_i^ν are constants and the a_i^{III} vanish, it follows* that the scalars $\kappa_{\nu 1}(t), \kappa_{\nu 2}(t)$ are independent of t if $\nu = 1, 2, 3$, and vanish if $\nu = 3$. Hence,

* By choosing subsequently $\nu = I, II, III$ and then writing 1, 2, 3 for I, II, III.

$$(16_1) \quad \kappa_{12} + \kappa_{21} = \text{const.}, \quad \kappa_{11} - \kappa_{22} = \text{Const.}; \quad (16_2) \quad \kappa_{31} \equiv 0, \quad \kappa_{32} \equiv 0.$$

On substituting (4₁)–(4₂) into the definition (6₂) of the κ_{pq} , one sees that (16₁)–(16₂) may be written as

$$(17_1) \quad r^3 s_1 s_2 = \text{const.}, \quad r^3 (s_1^2 - s_2^2) = \text{Const.}, \quad (r = r(t) > 0);$$

$$(17_2) \quad -2r's_2 + r(-s_2' + s_3 s_1) = 0, \quad 2r's_1 + r(s_1' + s_3 s_2) = 0,$$

respectively. And (17₁) means that $r^3 s_1^2$ and $r^3 s_2^2$ are independent of t , i.e., that there exist two constants c_1, c_2 for which

$$(18) \quad s_1 = c_1 r^{-\frac{1}{3}}, \quad s_2 = c_2 r^{-\frac{1}{3}}, \quad (r > 0).$$

Direct substitution shows that (17₂) reduces in virtue of (18) to

$$s_3 r c_1 - \frac{1}{2} r' c_2 = 0, \quad \frac{1}{2} r' c_1 + s_3 r c_2 = 0.$$

This is a pair of homogeneous linear equations which are satisfied by c_1, c_2 and have the determinant $s_3^2 r^2 + \frac{1}{4} r'^2$, where $r > 0$. Since this square sum cannot vanish unless both s_3, r' vanish, it follows that if at least one of the two constants c_1, c_2 does not vanish, then both functions s_3, r' vanish for every t . In other words, there must be satisfied at least one of the two pairs of conditions

$$(19_1) \quad c_1 = 0, \quad c_2 = 0; \quad (19_2) \quad s_3(t) \equiv 0, \quad r(t) = \text{const.}$$

In the case (19₁), both functions s_1, s_2 vanish, by (18), for every t . This means, by §72, that the rotation $\Omega(t)$ takes place about the ξ^{III} -axis of the inertial coordinate system for every t . But the ξ^{III} -axis was chosen to be such that every $\xi_i^{\text{III}} = \xi_i^{\text{III}}(t^0)$ vanishes. Hence, it is clear from (1) that every $\xi_i^{\text{III}}(t)$ vanishes for every t . In other words, every m_i moves in the $(\xi^{\text{I}}, \xi^{\text{II}})$ -plane of the inertial coordinate system. This proves (ii), §371 for the first of the two possible cases (19₁), (19₂).

In the case (19₂), one sees from (18) that all three functions s_1, s_2, s_3 are independent of t and the constant s_3 vanishes. Since all three s_i are constants, §75 shows that the rotation $\Omega = \Omega(t)$ takes place about an axis which has an invariable position with reference to the inertial coordinate system $\xi = (\xi^{\text{I}}, \xi^{\text{II}}, \xi^{\text{III}})$. Furthermore, since the constant s_3 vanishes, this fixed axis of rotation must lie within the $(\xi^{\text{I}}, \xi^{\text{II}})$ -plane (cf. the proofs in §71–§75). But the ξ^{III} -axis was chosen so that every $\xi_i^{\text{III}} = \xi_i^{\text{III}}(t^0)$ vanishes. Hence, it is seen from (1) that the rotating coordinate system $x = \Omega^{-1}\xi$, where $\Omega = \Omega(t)$, cannot actually rotate about a fixed axis contained in the $(\xi^{\text{I}}, \xi^{\text{II}})$ -

plane. Consequently, there is no rotation at all, i.e., $\Omega(t) = \text{const.}$ This means that the homographic solution under consideration is homothetic. Since it is clear from the definitions that a flat homothetic solution is a planar solution, (ii), §371 follows for the second of the two possible cases (19₁), (19₂).

This completes the proof of all the statements of §370 bis–§371.

§374 bis. One might think the preceding proof unnecessarily complicated; in fact, it seems to be plausible that (ii), §371 is a direct consequence of the homogeneity of the force function U , if one takes into account the conservation of the angular momentum and of the centre of mass (§316–§317).

Actually, such is not the case. In fact, it will be shown that if the attraction should be inversely proportional to the third, instead of the second, power of the distance, then (ii), §371 would be false even in the case of $n = 3$ bodies, although the ten integrals hold without change in this case also. It is not surprising that Lagrange considered as the principal achievement of his theory of the homographic solutions of the problem of $n = 3$ bodies the proof of the fact that every homographic solution is planar in the case $n = 3$ (in which case every solution is flat, of course).

Assuming that the attraction between the $n = 3$ bodies m_i is proportional to the third power of the distance, one has

$$(I) \quad m_i \xi_i'' = U_{\xi_i}, \quad m_i \eta_i'' = U_{\eta_i}, \quad m_i \zeta_i'' = U_{\zeta_i}, \quad (i = 1, 2, 3),$$

$$(II) \quad U = \frac{1}{2} \sum^* m_j m_k \{ (\xi_j - \xi_k)^2 + (\eta_j - \eta_k)^2 + (\zeta_j - \zeta_k)^2 \}^{-1},$$

where the scalars ξ_i, η_i, ζ_i denote “inertial” barycentric Cartesian coordinates of m_i , the summation (II) runs over the three cyclic permutations of $(j, k) = (1, 2)$, and the factor of proportionality, $\frac{1}{2}$, in (II) makes the choice of the units such that the force between two particles of mass 1 at distance 1 becomes 1. Choose the masses and the initial positions such that

$$(III) \quad m_1 = m_2;$$

$$(IV) \quad \xi_1^0 = -\xi_2^0 < 0 = \xi_3^0; \quad \eta_1^0 = \eta_2^0 < 0 < \eta_3^0; \quad \zeta_1^0 = \zeta_2^0 = \zeta_3^0 = 0,$$

where the superscripts ⁰ refer to $t = 0$. Since the coordinate system (ξ, η, ζ) is barycentric,

$$(V) \quad \sum_{i=1}^3 m_i \xi_i^0 = 0, \quad \sum_{i=1}^3 m_i \eta_i^0 = 0; \quad \text{hence,} \quad \sum_{i=1}^3 U_{\xi_i}^0 = 0, \quad \sum_{i=1}^3 U_{\eta_i}^0 = 0,$$

by (I). And (III)–(IV) are compatible with (V).

The content of (III)–(IV) is that the triangle formed by the three bodies at the date $t = 0$ is chosen as an isosceles triangle which lies in the (ξ, η) -plane and has two equal masses at its base; that the position of this base is chosen so as to be symmetric with respect to the η -axis; and that, if $t = 0$, the three m_i are ordered so that the increase of i determines the positive orientation of the (ξ, η) -plane. Thus, it is clear for reasons of symmetry that the components $U_{\xi_i}^0$, $U_{\eta_i}^0$, $U_{\zeta_i}^0$ of the force of attraction which acts on m_i when $t = 0$ are such that

$$(VI) \quad \begin{aligned} U_{\xi_1}^0 = -U_{\xi_2}^0 > 0 = U_{\xi_3}^0; \quad U_{\eta_1}^0 = U_{\eta_2}^0 > 0 > U_{\eta_3}^0; \\ U_{\zeta_1}^0 = U_{\zeta_2}^0 = U_{\zeta_3}^0 = 0; \end{aligned}$$

cf. (II), (III), (IV).

Put $a = -U_{\xi_1}^0 : m_1 \xi_1^0$ and $b = -U_{\eta_1}^0 : m_1 \eta_1^0$. Then $a > 0$ and $b > 0$, by (IV), (VI). Furthermore, (III), (IV), (VI) show that the relations

$$(VII_1) \quad U_{\xi_i}^0 = -am_i \xi_i^0, \quad (VII_2) \quad U_{\eta_i}^0 = -bm_i \eta_i^0$$

hold not only for $i = 1$ but for $i = 2$ also. Hence, it is clear from (V) that (VII₁)–(VII₂) hold for $i = 3$ as well. Finally, it is easily inferred from (II), (III), (IV) either by direct calculation or by an equivalent elementary vector consideration, that the relative magnitude of the two positive numbers a, b occurring in (VII₁)–(VII₂) depends on whether the side $m_1 m_3 = m_2 m_3$ of the isosceles triangle $m_1 m_2 m_3$ belonging to $t = 0$ is shorter than, equal to or longer than its base $m_1 m_2$, an equilateral triangle being characterized by $a = b$. Choose the initial position of m_3 so that

$$(VIII) \quad b > a, \quad (a > 0, b > 0).$$

It will be shown that the 9 initial velocities $\xi_1'^0, \dots, \zeta_3'^0$ may be chosen so that the solution of (I) which belongs to the 18 initial conditions $\xi_1^0, \dots, \zeta_3'^0$ becomes of the form

$$(IX) \quad \xi_i = \xi_i^0 r, \quad \eta_i = \eta_i^0 r \cos \omega, \quad \zeta_i = \zeta_i^0 r \sin \omega, \quad (i = 1, 2, 3),$$

where $r = r(t)$, $\omega = \omega(t)$ is a pair of suitably chosen functions which are not independent of t and which satisfy the initial conditions

$$(X) \quad r^0 = 1, \quad \omega^0 = 0 \quad (\text{cf. (IV), where } \zeta_i^0 = 0).$$

First, direct substitution of (IX), (II) into (I) shows that the $3 + 3 + 3$ conditions (I) for the 2 unknowns $r(t)$, $\omega(t)$ consist, on the one hand, of the 3 equations $m_i \xi_i^0 r'' = r^{-3} U_{\xi_i}^0$, which, in view of (VII₁), reduce to the 1 condition $r'' = -ar^{-3}$; and, on the other hand, of $3 + 3$ equations which, upon an application of the multipliers $\cos \omega$, $\sin \omega$ and $-\sin \omega$, $\cos \omega$, easily reduce to the 2 conditions $r'' - r\omega'^2 = -br^{-3}$, $r\omega'' + 2r'\omega' = 0$, if use is made of (VII₂) and of the fact that $U_{\xi_i}^0 = 0$, by (VI). But these $1 + 2$ conditions for the 2 functions $r(t)$, $\omega(t)$ are not independent. In fact, $r'' - r\omega'^2 = -br^{-3}$ becomes in virtue of $r'' = -ar^{-3}$ equivalent to $\omega' = (b - a)^{\frac{1}{2}} r^{-2}$, a condition in virtue of which $r\omega'' + 2r'\omega' = 0$ becomes an identity in t , since a, b are constants. Accordingly, (IX) is a solution of (I) if and only if $r(t)$ and $\omega(t)$ satisfy the pair of conditions

$$(XI) \quad r'' = -ar^{-3}, \quad \omega' = (b - a)^{\frac{1}{2}} r^{-2}.$$

It is readily verified that (XI) is satisfied by the functions

$$(XII) \quad \begin{aligned} r &\equiv r(t) = (1 + 2a^{\frac{1}{2}}t)^{\frac{1}{2}}, \\ \omega &\equiv \omega(t) = \frac{1}{2}a^{-\frac{1}{2}}(b - a)^{\frac{1}{2}} \log(1 + 2a^{\frac{1}{2}}t), \end{aligned}$$

which satisfy (X) also. And (VIII) shows that the constants $a^{-\frac{1}{2}}$, $(b - a)^{\frac{1}{2}}$ occurring in the solution (XII) of (XI) are real and such that $\omega(t) \neq \text{const.}$

It follows that the particular solution of (I) which is represented by (IX) and (XII) has the desired properties. For it is clear from (IX) that this solution of (I) is homographic, but not planar, since $\omega(t) \neq \text{const.}$; although the solution is flat, since $n = 3$.

It also follows that the result of §346 does not hold in case of the force function (II). In fact, it is clear from (IV) that the non-planar solution (IX) is such that the triangle formed by the three bodies is, at every t , an isosceles triangle which has the masses (III) at its base. Nevertheless, the angle $\omega(t)$ defined by (XII) is not constant; so that the fixed axis or plane of symmetry, established in §346 for the case of Newtonian gravitation, does not exist in the present case.

According to (XII), the functions $r(t)$, $\omega(t)$ are real on the interval $-\frac{1}{2}a^{-\frac{1}{2}} < t < +\infty$ and tend, as $t \rightarrow -\frac{1}{2}a^{-\frac{1}{2}} + 0$, to $\lim r = 0$, $\lim \omega = -\infty$. Hence, it is seen from (IX) that, as $t \rightarrow -\frac{1}{2}a^{-\frac{1}{2}} + 0$, the three bodies participate in a simultaneous collision in such a way that all three bodies move along non-planar spirals before colliding at the centre of mass. According to §335 and §326, a simultaneous

collision of the $n = 3$ bodies is impossible in case of a non-planar solution, if the attraction is Newtonian.

Homographic Solutions and Central Configurations

§375. The results collected in §370 bis–§371 and proved in §372–§374 contain a classification of all possible homographic solutions but leave open the question of the existence of such solutions. According to §369, such a solution, if any, is determined, on the one hand, by a pair of functions $r(t)$, $\Omega(t)$, and, on the other hand, by n initial position vectors ξ_i^0 .

In preparation for the treatment of the existence question, it will now be shown that the ξ_i^0 must be chosen so as to form a central configuration belonging to the given m_i . This, when compared with §355 and (1), §370, may be expressed also by saying that if a solution $\xi_i = \xi_i(t)$ belonging to n given m_i is homographic, then the m_i must form a central configuration at every t .

If the solution is planar, the inertial coordinate system ξ will always be chosen so that the paths lie in the plane $\xi^{III} = 0$, and $\phi' = \phi'(t) \geq 0$ will denote the angular velocity of the rotating plane (x^I, x^{II}) , where $x = \Omega^{-1}\xi$. If the solution is non-planar, let $\phi'(t)$ be defined by $\phi' \equiv 0$. Then, as shown at the end of §371, all formulae of §370–371 are valid in both cases. Thus, if the constants $m^0; h^0, C^0$ are defined by

$$(20_1) \quad m^0 = U^0/J^0;$$

$$(20_2) \quad h^0 = h/J^0; \quad (20_3) \quad C^0 = C/J^0, \quad (U^0 > 0, J^0 > 0),$$

then, from (13) and (8₂),

$$(21_1) \quad \frac{1}{2}(r'^2 + r^2\phi'^2) - m^0/r = h^0; \quad (21_2) \quad r^2\phi' = |C^0|.$$

Since (20₁)–(20₂) show also that (10) may be written as $r'^2 = -rr'' + m^0/r + 2h^0$, one sees from (21₁) that

$$(22) \quad r'' - r\phi'^2 = -m^0/r^2; \quad \text{while} \quad r\phi'' + 2r'\phi' = 0,$$

since $r\phi'' + 2r'\phi' \equiv (r^2\phi')'/r$ and $r^2\phi' = \text{const.}$, by (21₂).

Since a_i in (6₁) was defined by $m_i a_i = U_{\xi_i}^0$, one has $K\xi_i^0 = m_i^{-1}U_{\xi_i}^0$; and the third components of the 3-vectors ξ_i^0 , $U_{\xi_i}^0$ vanish in the planar case. On the other hand, K is, in view of (9) and (22), the diagonal 3-matrix formed by the diagonal elements $-m^0, -m^0, -m^0 + r^3\phi'^2$; and ϕ' vanishes in the non-planar case. Consequently, $-m^0\xi_i^0 = m_i^{-1}U_{\xi_i}^0$ in both the planar and non-planar cases. Thus, the con-

dition $U_{\xi_i} = \sigma m_i \xi_i$ of §355 for a central configuration is satisfied by $\sigma = -m^0$, if $t = t^0$. Since the initial date t^0 may be chosen arbitrarily, the proof is complete.

§376. Since $\phi' \equiv 0$ in the non-planar case, one can write the definition (1), §370 of a homographic solution $\xi_i = \xi_i(t)$ not only in the planar but also in the non-planar case in the form

$$(23) \quad \xi_i = r\Omega\xi_i^0, \text{ where } r = r(t); \Omega = \Omega(t) \equiv \begin{pmatrix} \cos \phi & -\sin \phi & 0 \\ \sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Hence, every homographic solution is determined by the n initial positions ξ_i^0 and by a pair of functions $r(t)$, $\phi(t)$ which are, by (2₁)-(2₂) and (21₂), subject to the trivial normalizations

$$(24_1) \quad r^0 = 1, \quad (r = r(t) > 0); \quad (24_2) \quad \phi^0 = 0; \quad (24_3) \quad \phi'^0 \geq 0.$$

This pair of functions may be described directly, as follows:

Interpret r , ϕ as polar coordinates in a Cartesian (x, y) -plane, and, choosing a fixed positive number m^0 arbitrarily, consider the dynamical problem with two degrees of freedom which is defined by the Lagrangian function

$$(25) \quad \begin{aligned} L &= \frac{1}{2}(x'^2 + y'^2) + m^0/(x^2 + y^2)^{\frac{1}{2}}; \quad \text{so that} \\ L &= \frac{1}{2}(r'^2 + r^2\phi'^2) + m^0/r. \end{aligned}$$

The Lagrangian equations $[L]_x = 0$, $[L]_y = 0$ are seen to be $x'' = -m^0x/r^3$, $y'' = -m^0y/r^3$ and admit, therefore, besides the energy integral $\frac{1}{2}(x'^2 + y'^2) - m^0/r = \text{Const.}$, the integral $xy' - yx' = \text{const.}$ But $x = r \cos \phi$, $y = r \sin \phi$; so that these integrals are identical with the relations (21₁)-(21₂), where $h^0 = \text{Const.}$, $|C^0| = \text{const.}$ On the other hand, it is seen from the second representation of L in (25), that the Lagrangian equations $[L]_x = 0$, $[L]_y = 0$, when expressed in terms of the polar coordinates r , ϕ in the form $[L]_r = 0$, $[L]_\phi = 0$ (cf. §95), are precisely the equations (22). Finally, comparison of (25) with §241 shows that the motion in the (x, y) -plane is that of a particle of unit mass in a static field of force; this field of force being generated by an ideal body which has the mass m^0 , rests at the origin $(x, y) = (0, 0)$, and attracts the moving particle according to Newton's law of gravitation, without being attracted by this particle. In other words, the problem of the determination of the pair of functions $r(t)$, $\phi(t)$, being identical with

the problem of integration of the Lagrangian equations (22) or $[L]_x = 0$, $[L]_y = 0$, is identical with the problem discussed in §241–§273, if one chooses $m^0 = 1$.

§377. It is now easy to construct homographic solutions. In fact, it will be shown that a solution $\xi_i = \xi_i(t)$ of the problem of n bodies, with given values m_i of the masses, is homographic if and only if there exist two functions $r(t)$, $\phi(t)$ and n initial position vectors ξ_i^0 by means of which $\xi_1(t), \dots, \xi_n(t)$ are representable in the form (23), (24₁)–(24₂), where $r = r(t)$, $\phi = \phi(t)$ may be chosen as any solution of the Lagrangian equations (22) belonging to (25) and satisfying (24₁)–(24₂), while ξ_1^0, \dots, ξ_n^0 is any central configuration belonging to m_1, \dots, m_n . It is understood that the constant m^0 occurring in (22) has to be defined in terms of the m_i and ξ_i^0 by placing, in accordance with (20₁)–(20₂),

$$(26_1) \quad m^0 = J^0/U^0;$$

$$(26_2) \quad J^0 = \sum m_i |\xi_i^0|^2; \quad (26_3) \quad U^0 = \sum^* m_j m_k / |\xi_j^0 - \xi_k^0|.$$

It has already been proved in §376 that $r(t)$, $\phi(t)$ must satisfy (22), and in §375, that ξ_1^0, \dots, ξ_n^0 must form a central configuration belonging to the m_i , if the solution $\xi_i = \xi_i(t)$ is homographic.

In order to prove that these necessary conditions are sufficient as well, one has only to show that, when they are satisfied, the functions $\xi_1(t), \dots, \xi_n(t)$ defined by (23) are solutions of the problem of n bodies. For it is clear that if the $\xi_i(t)$ are of the form (23), then $\xi_i = \xi_i(t)$ represents either an homographic solution or no solution at all. But the condition imposed on the ξ_i^0 is that there exists a scalar σ for which $U_{\xi_i}^0 = \sigma m_i \xi_i^0$, in which case σ has necessarily the value $\sigma = -U^0/J^0$ (cf. §355); so that $U_{\xi_i}^0 = -m^0 m_i \xi_i^0$, by (26₁). Since (3₃) is implied by (23), it follows that the equations of motion, $m_i \xi_i'' = U_{\xi_i}$, reduce to $r^2 \Omega^{-1} \xi_i'' = -m^0 \xi_i^0$. Consequently, one has only to show that $r^2 \Omega^{-1} \xi_i'' = -m^0 \xi_i^0$ is an identity in t in virtue of (23) and (22).

To this end, let $r = r(t)$, $\phi = \phi(t)$ be any given pair of functions which have continuous second derivatives. Let $\Omega = \Omega(t)$ be defined in terms of $\phi = \phi(t)$ as that 3-matrix Ω which is of the type occurring in (23). Finally, let a 3-matrix $K = K(t)$ be defined in terms of $r = r(t)$, $\phi = \phi(t)$ by means of (9). Then it is easily shown by straightforward differentiations and matrix multiplications, that the product of $r^2 \Omega^{-1}$ and $(r\Omega)''$ is identical with K . Since (23) implies

that $\xi_i'' = (r\Omega)''\xi_i^0$, it follows that $r^2\Omega^{-1}\xi_i'' = K\xi_i^0$ is an identity in t . Consequently, one has only to show that $K\xi_i^0 = -m^0\xi_i^0$ is an identity in virtue of (22). But this has already been verified at the end of §375, since the constant vector $m_i^{-1}U_{\xi_i}^0$ considered there was seen to be identical with the constant vector $-m^0\xi_i^0$. This completes the proof of the criterion announced at the beginning of this article.

§377 bis. It is clear from the proof given in §377, that in order that a solution of the problem of n bodies be homographic, it is not only necessary (§375) but also sufficient that the m_i form the same* central configuration for every t .

§378. The variety of all homographic solutions belonging to n given m_i may now be enumerated, as follows:

Choose an arbitrary central configuration ξ_1^0, \dots, ξ_n^0 belonging to m_1, \dots, m_n and define three positive numbers by (26₁)–(26₃). Since $\beta\xi_1^0, \dots, \beta\xi_n^0$ determine, for every $\beta > 0$, the same central configuration as ξ_1^0, \dots, ξ_n^0 (cf. the end of §355), and belong, by (26₂)–(26₃), to β^2J^0 and $\beta^{-1}U^0$, one sees from (26₁) that the given central configuration may be assumed to be such as to satisfy the condition $m^0 = 1$, mentioned at the end of §376. Then the solutions

$$(27) \quad x = x(t), \quad y = y(t), \quad \text{i.e.,} \quad r = r(t), \quad \phi = \phi(t),$$

of the Lagrangian equations belonging to (25) are exactly those discussed in §241. Choose the four initial values $r^0, \phi^0; r'^0, \phi'^0$, assigned to these Lagrangian equations (22) at an initial $t = t^0$, in such a way that r^0, ϕ^0 and the sign of ϕ'^0 are given by (24₁)–(24₃). Then, on applying at $t = t^0$ the integrals (21₁)–(21₂) of (22), where $m^0 = 1$, one sees that r'^0 and ϕ'^0 follow from

$$(28_1) \quad \frac{1}{2}(r'^0)^2 + \frac{1}{2}(\phi'^0)^2 - 1 = h^0; \quad (28_2) \quad \phi'^0 = |C^0|,$$

where $h^0, |C^0|$ are, in the sense of §241, the energy and the angular momentum of the solution path (27) in an (x, y) -plane; so that the constants $h^0, |C^0|$ defined by (28₁)–(28₂) are identical with the constants which in §241 were denoted by h, c , where c may be chosen to satisfy $c = |c| \geq 0$ without loss of generality (cf. §242). Since the

* This is meant in the sense defined at the end of §355. Notice that if there exists a continuum of distinct central configurations for n given m_i (cf. §365), it might occur that these m_i form a central configuration for every t in a suitable solution which is not homographic, since the central configuration might then vary with t .

initial values $r'^0, \phi'^0 \geq 0$ may be chosen arbitrarily, it is clear from (28₁)–(28₂) that all three cases $h^0 \geq 0$ are possible in both cases $|C^0| \geq 0$.

It is seen from §242 that if (28₂) is chosen to be 0, then, and only then, is the path (27) rectilinear in the (x, y)-plane (leading to the rectilinear limiting forms of hyperbolic, parabolic and elliptic motions according as $h^0 \geq 0$). If, on the other hand, (28₂) is chosen to be distinct from 0, then (27) is, again by §241, a branch of an hyperbola, a parabola or an ellipse according as the constant (28₁) is chosen to be ≥ 0 . Finally, §377 assures that, in all six cases $|C^0| \geq 0, h^0 \geq 0$, substitution of (27) into (23) defines an homographic solution $\xi_i = \xi_i(t)$ of the problem $m_i \xi_i'' = U_{\xi_i}$ of n bodies m_i . According to (20₂), the energy constant h of this solution is of the same sign as the energy constant (28₁); while comparison of (20₃) with (I), §370 bis shows that the solution is homothetic if and only if the momentum constant (28₂) is chosen to be 0.

It follows, in particular, that there exist for every central configuration of the m_i homothetic solutions of arbitrary energy $h \geq 0$. Notice that in order that there exists for every m_i a line l_i which has an invariable position with reference to the inertial coordinate system ξ and contains m_i for every t , i.e., in order that the homographic solution (23) be homothetic, it is, by (21₂) and (I), §370 bis, necessary and sufficient that the path (27) in an (x, y)-plane be rectilinear. Notice, however, that in order that the latter path be rectilinear, it is not necessary (though, of course, sufficient) that paths of the individual m_i be rectilinear. In fact, one can choose the momentum constant (28₂) to be distinct from 0 also when the given central configuration ξ_1^0, \dots, ξ_n^0 is collinear in the sense of §355. The simplest instance of this situation follows from §378 bis below.

On the other hand, the solution path (27) will reach the origin $r = 0$ of the (x, y)-plane at some t if and only if (28₂) is chosen to be 0. It follows, therefore, from (23) that in case of an homographic solution the non-existence of an invariable plane, i.e., $C = 0$, is not only necessary (§335) but also sufficient for a simultaneous collision of all n bodies. The deep result of §363–§364 is, of course, trivial in this rather particular case of a simultaneous collision. Also the problem mentioned in §368 does not arise in this case.

Finally, let the constant (28₂) be chosen distinct from 0. Then $\phi' \neq 0$, by (21₂); so that, by §371, the homographic solution (23) is necessarily planar. In particular, the given central configuration ξ_1^0, \dots, ξ_n^0 is then flat in the sense of §355, collinear configurations

being not excluded. Since $|C^0| \neq 0$, it follows from (4), §241 that the path (27) in an (x, y) -plane is an ellipse or an hyperbola with the major axis $-1/h^0 = 2a \geq 0$ and the eccentricity $(1 + 2h^0|C^0|^2) = e \leq 1$ ($\neq e$) according as $h^0 \leq 0$; and that it is a parabola with the parameter $|C^0|^2 = p \neq 0$ if $h^0 = 0$; finally, that it has $(x, y) = (0, 0)$ as a focus in all three cases. Since $r'^0, \phi'^0 (> 0)$ in (28₁)–(28₂) may be chosen arbitrarily, the same holds for $h^0, |C^0| (> 0)$, and, therefore, also for $2a (\neq 0)$, $e (\neq 1)$ or $p (\neq 0)$, where $a \geq 0$, $e \geq 0$ or $p > 0$. And substitution of (27) into (23) shows that the n bodies move, in all three cases $h^0 \geq 0$, along n co-planar and similar conics in the (ξ^I, ξ^{II}) -plane of the barycentric inertial coordinate system whose origin is a common focus of all n conics. The figure illustrates the situation for $h^0 < 0$ in the case of an equilateral central configuration belonging to $n = 3$ masses m_i (cf. §367).

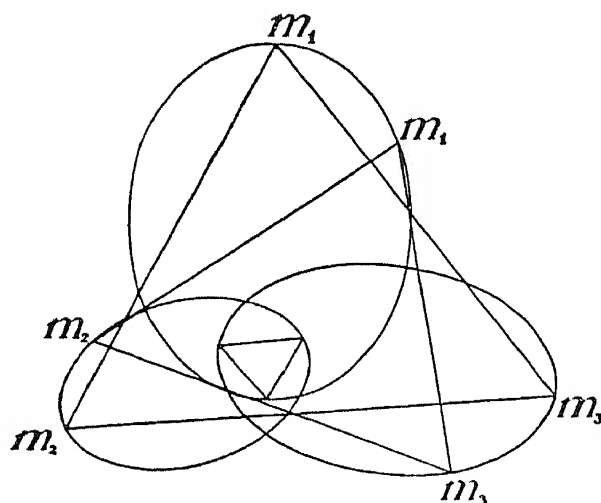


FIG. 13

The n conics are circles if and only if $e = 0$, i.e., $1 + h^0|C^0|^2 = 0$. This condition is, by (28₁)–(28₂), equivalent to $r'^0 = 0$ or, since t^0 may be chosen arbitrarily, to $r'(t) \equiv 0$. In other words, a planar homographic solution (23) satisfies the defining condition $r(t) = \text{const.}$ of a solution of relative equilibrium if and only if the n paths in the inertial (ξ^I, ξ^{II}) -plane are concentric circles about the centre of mass $\xi = 0$. But, as seen above, h^0 and $|C^0| > 0$ may be chosen arbitrarily in case of any homographic non-homothetic solution; so that $1 + h^0|C^0|^2 = 0$ can be satisfied in case of any flat central configuration. Furthermore, all solutions of relative equilibrium are planar, by (II), §370 bis. Consequently, there exists for every flat, and for no non-flat, central configuration a solution of relative equilibrium.

§378 bis. The above results apply, in particular, to any solution of the problem of $n = 2$ bodies. In fact, the configuration formed by two arbitrary masses is, by §359, always a central configuration. It follows, therefore, from §377 bis that every solution $\xi_1 = \xi_1(t)$, $\xi_2 = \xi_2(t)$ of the problem of $n = 2$ bodies is homothetic. Actually, this is implied by the barycentric condition $m_1\xi_1 + m_2\xi_2 = 0$ also. That every solution of the problem of $n = 2$ bodies is planar, is implied not only by §207 and (13), §343 but also by §329 (and, of course, by (ii), §371).

§379. Without any reference to homothetic solutions, consider the planar problem of n bodies; so that the Lagrangian function $L = T + U$ is given by

$$(29_1) \quad L = \frac{1}{2} \sum m_i (\xi_i'^2 + \eta_i'^2) + \sum^* m_j m_k / \rho_{jk};$$

$$(29_2) \quad \rho_{jk}^2 = (\xi_j - \xi_k)^2 + (\eta_j - \eta_k)^2,$$

where the scalars ξ, η denote, for simplicity, the components ξ^I, ξ^{II} of the 3-vector $(\xi^I, \xi^{II}, \xi^{III}) \equiv (\xi^I, \xi^{II}, 0)$ in a barycentric inertial coordinate system. Besides the barycentric inertial coordinate plane (ξ, η) , consider a barycentric non-inertial plane (x, y) which rotates about the common origin with some given constant angular velocity, say ω ; so that

$$(30) \quad \xi_i = x_i \cos \omega t - y_i \sin \omega t, \quad \eta_i = x_i \sin \omega t + y_i \cos \omega t,$$

if the origin of the t -axis is chosen so that $(x, y) = (\xi, \eta)$ at $t = 0$. It is easily verified from (30) that

$$(31_1) \quad \xi_i'^2 + \eta_i'^2 = (x_i' - \omega y_i)^2 + (y_i' + \omega x_i)^2;$$

$$(31_2) \quad \rho_{jk}^2 = (x_j - x_k)^2 + (y_j - y_k)^2.$$

On substituting (31₁)–(31₂) into (29₁) and then carrying out the Lagrangian differentiations, one readily finds that the equations of motion, $[L]_{x_i} = 0$ and $[L]_{y_i} = 0$, in terms of the coordinates of the rotating plane (x, y) are

$$(32) \quad m_i(x_i'' - 2\omega y_i' - \omega^2 x_i) = U_{x_i}, \quad m_i(y_i'' + 2\omega x_i' - \omega^2 y_i) = U_{y_i},$$

where $U = \sum^* m_j m_k / \rho_{jk}$ is thought of as expressed by means of (31₂).

Since all this holds for any $\omega = \text{const.}$ in (30), it follows from (II), §370 bis that a solution $\xi_i = \xi_i(t)$, $\eta_i = \eta_i(t)$ of relative equilibrium is characterized by the existence of a suitable value of $\omega = \text{const.}$

such that the system (32) has, for this particular ω , a solution of the form

$$(33_1) \quad x_i(t) \equiv x_i^0, \quad y_i(t) \equiv y_i^0; \quad (33_2) \quad \sum m_i x_i^0 = 0, \quad \sum m_i y_i^0 = 0,$$

where x_i^0, y_i^0 are suitable scalar constants satisfying the barycentric conditions (33₂).

Substitution of (33₁) into (32) gives

$$(34) \quad \omega^2 x_i^0 = \sum_{k=1}^n {}'m_k \frac{x_i^0 - x_k^0}{{}^0\rho_{ik}^3}, \quad \omega^2 y_i^0 = \sum_{k=1}^n {}'m_k \frac{y_i^0 - y_k^0}{{}^0\rho_{ik}^3}$$

$${}^0\rho_{ik} = \left\{ (x_i^0 - x_k^0)^2 + (y_i^0 - y_k^0)^2 \right\}^{\frac{1}{2}},$$

where the dashes ' indicate that $k \neq i$. It follows, therefore, from the last remark of §378 that the problem of determining all sets of $2n + 1$ constants $x_i^0, y_i^0; \omega$ which satisfy the $2n + 2$ conditions (34), (33₂) is equivalent to the problem of enumerating all flat central configurations belonging to the given m_i (cf. §360).

It is clear from (34), (33₂) that if (33₁) is a solution of relative equilibrium belonging to given m_i and to the angular velocity ω , then $x_i(t) = \rho x_i^0, y_i(t) = \rho y_i^0$ is, for any positive number ρ , a solution belonging to the same m_i and to the angular velocity $\rho^{-\frac{1}{2}}\omega$ (this agrees with the remarks made at the end of §315). Incidentally, this arbitrary change of the linear dimensions, together with the possible passage from t to $\pm t + \text{const.}$, exhausts all solutions of relative equilibrium belonging to one and the same central configuration of the m_i (cf. the end of §355). In fact, the end of §378 shows that one has to satisfy the condition $1 + 2h^0|C^0|^2 = 0$; so that the ratio $h^0:|C^0|^{-2}$ is uniquely determined. The sign of ω remains, of course, undetermined, since the passage from ω to $-\omega$ is, in view of (30), equivalent to the admissible passage from t to $-t$.

§380. As an illustration, the angular velocity of the solutions of relative equilibrium of the problem of $n = 3$ bodies will now be computed.

In the collinear case of n bodies, one can choose the x -axis of the rotating coordinate system (x, y) so that all $y_i^0 = 0$, and, in addition, assume that $x_i^0 < x_{i+1}^0$. Then (34) reduces to

$$(35) \quad \omega^2 x_i^0 = \sum_{k=1}^{i-1} \frac{m_k}{(x_i^0 - x_k^0)^2} - \sum_{k=i+1}^n \frac{m_k}{(x_i^0 - x_k^0)^2}, \quad 0 = 0,$$

since $x_i^0 - x_k^0 = \pm |x_i^0 - x_k^0| \equiv \pm {}^0\rho_{ik}$ according as $i \geq k$. It is understood that the first sum on the right of (35) is vacuous for $i = 1$, and the second for $i = n$. Thus, if $n = 3$, then (35) requires that $\omega^2 x_1^0, \omega^2 x_2^0, \omega^2 x_3^0$ be equal to

$$-\frac{m_2}{{}^0\rho_{12}^2} - \frac{m_3}{{}^0\rho_{13}^2}, \quad -\frac{m_3}{{}^0\rho_{23}^2} + \frac{m_1}{{}^0\rho_{12}^2}, \quad +\frac{m_1}{{}^0\rho_{13}^2} + \frac{m_2}{{}^0\rho_{23}^2},$$

respectively. If one forms the two linear combinations $\omega^2(x_2^0 - x_1^0) = \dots$, $\omega^2(x_3^0 - x_2^0) = \dots$ of these three conditions and observes that $x_2^0 - x_1^0 = {}^0\rho_{21}$, $x_3^0 - x_2^0 = {}^0\rho_{23}$; ${}^0\rho_{13} = {}^0\rho_{12} + {}^0\rho_{23}$, it follows that it is sufficient to determine three positive numbers ${}^0\rho_{12}, {}^0\rho_{23}; \omega^2$ satisfying the two conditions

$$(36) \quad \begin{aligned} {}^0\rho_{12}\omega^2 &= \frac{m_1 + m_2}{{}^0\rho_{12}^2} + \frac{m_3}{({}^0\rho_{12} + {}^0\rho_{23})^2} - \frac{m_3}{{}^0\rho_{23}^2}, \\ {}^0\rho_{23}\omega^2 &= \frac{m_3 + m_2}{{}^0\rho_{23}^2} + \frac{m_1}{({}^0\rho_{12} + {}^0\rho_{23})^2} - \frac{m_1}{{}^0\rho_{12}^2}. \end{aligned}$$

In fact, if ${}^0\rho_{21}$ and ${}^0\rho_{23}$ are known, then x_1^0, x_2^0, x_3^0 follow uniquely from the barycentric condition $\sum m_i x_i^0 = 0$.

On defining a 2-matrix (σ_{pq}) by

$$(37) \quad \begin{aligned} &\begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} \\ &= \begin{pmatrix} \omega^2 - \frac{m_1 + m_2}{{}^0\rho_{12}^3} - \frac{m_3}{({}^0\rho_{12} + {}^0\rho_{23})^3} & \frac{m_3}{({}^0\rho_{12} + {}^0\rho_{23})^3} - \frac{m_3}{{}^0\rho_{23}^3} \\ \frac{m_1}{({}^0\rho_{12} + {}^0\rho_{23})^3} - \frac{m_1}{{}^0\rho_{12}^3} & \omega^2 - \frac{m_3 + m_2}{{}^0\rho_{23}^3} - \frac{m_1}{({}^0\rho_{12} + {}^0\rho_{23})^3} \end{pmatrix}, \end{aligned}$$

one can write (36) in the form

$${}^0\rho_{12}\sigma_{11} = {}^0\rho_{23}\sigma_{12}, \quad {}^0\rho_{12}\sigma_{21} = {}^0\rho_{23}\sigma_{22}.$$

Since ${}^0\rho_{12}, {}^0\rho_{23}$ are positive, this implies not only that the determinant of (37) vanishes, but also that σ_{11}, σ_{22} are of the same sign as σ_{12}, σ_{21} , respectively. But σ_{12}, σ_{21} are negative, since in their definition (37) the respective factors of m_3, m_1 are positive. Hence,

$$(38_1) \quad \sigma_{pq} < 0 \quad (p = 1, 2; q = 1, 2); \quad (38_2) \quad \sigma_{11}\sigma_{22} - \sigma_{12}\sigma_{21} = 0.$$

Finally, on placing $\rho = {}^0\rho_{12} + {}^0\rho_{23}$ and $\lambda = {}^0\rho_{12} : {}^0\rho_{23}$ and then expressing the determinant of (37) in terms of ρ and λ , one readily finds that (38₂) may be written in the form

$$(39) \quad \omega^2 \rho^3 = m_1 + m_3 + m_2(1 + \lambda)^2(1 + \lambda^{-2}),$$

$\lambda = \lambda(m_1, m_2, m_3)$ being the unique positive root of (11), §358. Actually, (39) follows also without the use of the quintic equation (11), §358, if one adds the two relations (36) and observes that ${}^0\rho_{12}\omega^2 + {}^0\rho_{23}\omega^2 = \rho\omega^2$, while ${}^0\rho_{12} = \rho/(1 + \lambda)$, ${}^0\rho_{23} = \rho\lambda/(1 + \lambda)$.

In the remaining case of a solution of relative equilibrium of the problem of $n = 3$ bodies, the configuration is, by §359, an equilateral triangle. Hence, it is easily verified from (33₂) and (34), where $n = 3$, that

$$(40) \quad \omega^2 \rho^3 = m_1 + m_2 + m_3,$$

if ρ denotes the common value of the three sides ${}^0\rho_{ik}$.

In both cases (39), (40), the angular velocity $\pm \omega$ is seen to be proportional to the $-\frac{3}{2}$ -th power of the linear dimensions (cf. §379).

§381. It is clear that a solution (30) of relative equilibrium is characterized by the fact that (33₁), when combined with its consequence $x'_i(t) \equiv 0$, $y'_i(t) \equiv 0$, represents for (32) a solution which is an equilibrium solution in the sense of §83. It follows, therefore, from §89 that if $u_i, v_i; u'_i, v'_i$ denote the displacements (§86) of $x_i^0, y_i^0; 0, 0$ with reference to (32), then the system of the corresponding Jacobi equations (§86) has constant coefficients and is, therefore, of the form

$$(41) \quad z'_j = \sum_{l=1}^{4n} a_{jl} z_l, \quad (j = 1, \dots, 4n),$$

where $A = (a_{jl})$ is a constant $4n$ -matrix and z_1, \dots, z_{4n} denote the $4n$ displacements u_i, v_i, u'_i, v'_i ($i = 1, \dots, n$).

The values of the constant coefficients a_{jl} may be obtained by observing that (41) is the linear Lagrangian system

$$(42) \quad [L]_{u_i} = 0, \quad [L]_{v_i} = 0, \quad (i = 1, \dots, 2n)$$

whose Lagrangian function is a quadratic form with constant coefficients. In fact, application of the rule of §101 to the present case (29₁)-(33₁) shows that

$$(43) \quad L = \frac{1}{2} \sum m_i \{ (u'_i - \omega v_i)^2 + (v'_i + \omega u_i)^2 \} \\ + \frac{1}{2} \sum \sum \{ \alpha_{ik} u_i u_k + 2\beta_{ik} u_i v_k + \gamma_{ik} v_i v_k \},$$

where $\alpha_{ik} = U_{x_i x_k}^0$, $\beta_{ik} = U_{x_i y_k}^0$, $\gamma_{ik} = U_{y_i y_k}^0$ are constants, obtained by substituting the constants (33₁) into the second partial derivatives of $U = U(x_1, \dots, y_n) \equiv \sum^* m_i m_k / \rho_{ik}$. These partial differentiations and substitutions, which supply the coefficient matrix $A = (a_{jl})$ of (41), are, of course, quite tiresome in every case.

If A is computed, the equation $\det(sE - A) = 0$ of §89, which is now of degree $4n$, determines the characteristic exponents s . The discussion of the question, whether or not all $4n$ characteristic exponents s are of the stable type in the sense of §89, is still more tiresome than the calculation of the coefficient matrix A . In fact, one has to decide whether or not every root of the equation $\det(sE - A) = 0$ is purely imaginary (incl. 0), where E is the unit $4n$ -matrix.

§382. Let, in particular, $n = 3$; so that one has to deal with the two cases discussed in §380. Choose the unit of length so that $\rho = 1$ in both cases (39), (40), and put, for abbreviation,

$$(44_I) \quad \nu^2 = -\sigma_{12} - \sigma_{21};$$

$$(44_{II}) \quad \nu^2 = \frac{2.7}{4}(m_1 m_2 + m_2 m_3 + m_3 m_1) / (m_1 + m_2 + m_3)^2,$$

according as the configuration is collinear or equilateral, the σ_{pq} being defined by (37) in the first case and undefined in the second case. It is clear from (38₁) and (44_I)–(44_{II}) that ν^2 is positive in both cases.

On carrying out the elementary calculations assigned by §381, one finds that the equation $\det(sE - A) = 0$ of degree $4n = 12$ has, in either case, eight trivial roots of the stable type $s = \pm \tau\sqrt{-1}$, where τ assumes only the values $\tau = \omega$ and $\tau = 0$; and that the remaining four characteristic exponents s are the roots of

$$(45_I) \quad s^4 + (\omega^2 - \nu^2)s^2 - (2\nu^4 + 3\nu^2\omega^2) = 0;$$

$$(45_{II}) \quad s^4 + \omega^2 s^2 + \nu^2 \omega^4 = 0$$

according as the configuration is collinear or equilateral.

It follows that the answer to the question, whether or not all characteristic exponents are of the stable type in the sense of §89, is quite different in the two cases, since

(I) in the collinear case, one cannot choose the values of the three masses m_i such that all characteristic exponents become of the stable type; while

(II) in the equilateral case, all characteristic exponents are of the stable type or some of them are not, according as one or none of the three masses represents, respectively, more than $100(\frac{1}{3} + \frac{4}{9}\sqrt{2})$ percent or more than $100(\frac{1}{2} + \frac{1}{18}\sqrt{69})$ percent of the total mass,

m_1 , m_2 , m_3 (these limiting percentages are very high, somewhat higher than 96%; and both are below 97%).

In fact, (45_I) is a real quadratic equation in s^2 , with a negative constant term; so that one of the two roots s^2 of (45_I) is negative, the other positive. Hence, while one pair of the four roots s of (45_I) is purely imaginary, the other pair consists of a positive and of a negative number. Accordingly, two of the characteristic exponents s are not of the stable type, no matter what are the three given values m_i of the masses.

On the other hand, the two roots of the quadratic equation (45_{II}) for s^2 are given by $s^2 = \frac{1}{2} \{ (-1 \pm (1 - 4\nu^2)^{\frac{1}{2}}) \} \omega^2$. Hence, the four roots s of (45_{II}) are purely imaginary, distinct, non-vanishing numbers whenever $4\nu^2 < 1$, but they coincide pairwise as $4\nu^2 \rightarrow 1$ and become for $4\nu^2 > 1$ of the form $\pm \alpha \pm \beta\lambda = 1$, where α , β is a fixed pair of positive numbers. Accordingly, all characteristic exponents are of the stable type if and only if $4\nu^2 < 1$. And it is easily verified from (44_{II}) that $4\nu^2 < 1$ is equivalent to the percentual condition stated above.

§382 bis. As another instance, consider (cf. (iii), §360) the central configuration formed, on the one hand, by $n - 1$ bodies m_1, \dots, m_{n-1} which are placed at the corners of a regular $(n - 1)$ -gon and have the same mass, say m , and, on the other hand, of an n -th body which is placed at the mid-point of this $(n - 1)$ -gon and has the mass $m_n = 1$; so that the total mass is $\sum m_i = m(n - 1) + 1$. Maxwell found in his theory of the rings of Saturn* that the solution of relative equilibrium belonging to this central configuration has characteristic exponents which are all of the stable type at least as long as $m < 2/n^2$. Notice that this requires that $nm > 0$ when m_n is fixed and $n \rightarrow \infty$; so that the total mass, $m(n - 1)$, of the $n - 1$ bodies which form the "ring" is restricted to be the smaller the larger is n .

Elimination of the Linear Momentum

§383. The invariant relation $\sum m_i \xi_i = 0$ of the barycentric equations of motion $m_i \xi_i'' = U_{\xi_i}$ was, in §341, eliminated by introduction of the $n - 1$ heliocentric position vectors x_i , referred to m_n as Sun; so that

* Actually, Maxwell's calculation is arranged in such a way that m_n is assumed to have a fixed position; so that the action of the "ring" (m_1, \dots, m_{n-1}) on the "Saturn" m_n is neglected.

$$(1_1) \quad x_j = \xi_j - \xi_n;$$

$$(1_2) \quad \rho_{jk} = |x_j - x_k|, \quad \rho_{jn} = |x_j| \quad (j, k = 1, \dots, n-1).$$

The corresponding representation of $L = T + U$ was found to be

$$(2_1) \quad T = \frac{1}{2} \sum^0 m_j x_j'^2 - \frac{1}{2} (\sum^0 m_j x_j')^2 / \mu;$$

$$(2_2) \quad U = \sum^* m_j m_k / \rho_{jk} + m_n \sum^0 m_j / \rho_{jn},$$

(cf. (5)-(6), §341), where $\mu = \sum m_j$ and

$$(3_1) \quad \sum = \sum_{i=1}^n, \quad \sum^0 = \sum_{j=1}^{n-1};$$

$$(3_2) \quad \sum^* = \sum_{1 \leq j < k \leq n}, \quad \sum^+ = \sum_{1 \leq j < k \leq n-1} \quad (= \sum^* - \sum^0),$$

by (4₁)-(4₂), §341. Similarly, $J = \sum m_i \xi_i^2$ and $\sum m_i \xi_i \times \xi_i' = C$ became

$$(4_1) \quad J = \sum^0 m_j x_j^2 - (\sum^0 m_j x_j)^2 / \mu;$$

$$(4_2) \quad \sum^0 m_j (x_j \times x_j') - (\sum^0 m_j x_j) \times (\sum^0 m_j x_j') / \mu = C$$

(cf. (10₂), §342 and (9₁), §341 bis). Finally, the explicit form of the Lagrangian equations $[L]_{x_j} = 0$ is, by (11₁)-(11₂), §342,

$$(5_1) \quad x_j'' + (m_n + m_j) \frac{x_j}{|x_j|^3} = {}^i\Omega_{x_j};$$

$$(5_2) \quad {}^i\Omega = \sum_{k=1}^{n-1} {}'m_k \left(\frac{1}{|x_k - x_j|} - \frac{x_k \cdot x_j}{|x_k|^3} \right),$$

where $j = 1, \dots, n-1$, and the prime ' in (5₂) means that $k \neq j$.

The Hamiltonian form of these equations will now be determined. To this end, notice first that (6₁), §341, where $T = T(x_1', \dots, x_{n-1}', x_n')$, was obtained from $\frac{1}{2} \sum m_i \xi_i'^2$ by a non-singular linear transformation and is, therefore, a positive definite form in $x_1', \dots, x_{n-1}', x_n'$. Hence, it is clear from (6₁)-(6₂), §341 that the quadratic form (2₁), of the present article is positive definite in x_1', \dots, x_{n-1}' . In other words, (2₁), §155 is satisfied, and so §158 is applicable to the Lagrangian function $L = T + U$ defined by the formulae (2₁)-(2₂) of the present article. Accordingly, if y_1, \dots, y_{n-1} denote the 3-vectors whose components are canonically conjugate to the components of the 3-vectors x_1, \dots, x_{n-1} of heliocentric coordinates, then the

Hamiltonian function belonging to the Lagrangian function L . $T + U$ is obtained by expressing T in $H = T + U$ in terms of $y_j = L_{x'_j}$.

In order to carry out this calculation, notice first that

$$(6_1) \quad m_j^{-1}y_j = x'_j = \mu^{-1}\sum^0 m_k x_k; \quad (6_2) \quad m_j x'_j = y_j + m_n^{-1}m \sum^0 y_k.$$

In fact, $y_j = L_{x'_j} = T_{x'_j} + U_{x'_j} = T_{x'_j}$; so that (6₁) is clear from (2₁). On the other hand, (6₂) is the inverse of the linear transformation (6₁) of x'_1, \dots, x'_{n-1} into y_1, \dots, y_{n-1} . This is seen by first calculating $\sum^0 y_j$ from (6₁), and then observing that $\mu = \sum m_i$ may be written, in view of (3₁), as $\mu = m_n + \sum^0 m_j$; so that (6₂) is an identity in virtue of (6₁). And substitution of (6₂) into (2₁), (4₂) shows that

$$(7_1) \quad T = \frac{1}{2} \sum^0 m_j^{-1} y_j^2 + \frac{1}{2} (\sum^0 y_j)^2 / m_n; \quad (7_2) \quad \sum^0 x_j^2 \approx y_j^2 = C.$$

§384. Accordingly, the Hamiltonian form of the heliocentric Lagrangian equations (5₁) is

$$(8) \quad y'_j = -H_{y_j}, \quad x'_j = H_{x_j}, \quad \text{where} \\ H = T + U \text{ is given by (7₁), (2₂), (1₂).$$

On the other hand, the Hamiltonian form of the barycentric inertial Lagrangian equations $m_i \xi_i'' = U_{\xi_i}$ is, by §320,

$$(9_1) \quad \eta'_i = -H_{\xi_i}, \quad \xi'_i = H_{\eta_i}, \quad \text{with } H = \frac{1}{2} \sum m_i \xi_i'^2 = U(\xi_i);$$

$$(9_2) \quad \eta_i = m_i \xi'_i,$$

where (9₂) is implied by (9₁); so that, since $\sum m_i \xi_i = 0$ and $\sum m_i \xi_i \times \xi'_i = C$,

$$(10_1) \quad \sum m_i \xi_i = 0; \quad (10_2) \quad \sum \eta_i = 0; \quad (10_3) \quad \sum \xi_i \approx \eta_i = C.$$

Notice that j and i in (8) and (9₁) run from 1 to $n-1$ and to n , respectively, while (9₁) possesses the invariant system (10₁)–(10₃). The object of the passage from the system (9₁) with $3n$ degrees of freedom to the system (8) with $3(n-1)$ degrees of freedom is precisely the elimination of the barycentric conditions (10₁)–(10₃).

Since $\sum m_i \xi'_i = 0$ and $\sum m_i = \mu$, it is clear from (1₁) and (3₁) that $\sum^0 m_j x'_j = -\mu \xi'_n$. It follows, therefore, from (6₁) that $m_j^{-1}y_j = x'_j + \xi'_n$. Hence, $\eta_j = y_j$, by (1₁) and (9₂); so that the heliocentric momenta y_j are identical with the barycentric inertial momenta η_j , where $j = 1, \dots, n-1$. On the other hand, the heliocentric coordinates x_j are for every j distinct from the barycen-

tric) inertial coordinates ξ_j (except when m_n happens to be situated at the centre of mass, $\xi = 0$, of all n bodies). Thus, it is clear that the connection between the heliocentric momenta and velocities, y_j and x'_j , cannot be the same as that between the barycentric momenta and velocities, η_j and ξ'_j .

Actually, (9₂) and (6₂) show that each of the inertial barycentric and none of the heliocentric momenta is identical with a constant multiple of the respective velocity. This fact, which is usually expressed by saying that while (9₁) is, the result (8) of the heliocentric elimination of the invariant system (10₁)–(10₂) is not, of the osculating type, makes a use of the reduced system (8) quite inconvenient for practical application to problems of the type exemplified by the solar system. Furthermore, the fact that the quadratic form, (2₁) or (7₁), which represents the kinetic energy in case of heliocentric coordinates is not diagonal, may become bothersome in theoretical investigations also (cf., in particular, §415–§420 below). For these reasons, the heliocentric coordinates x_j will now be replaced by certain of their linear combinations, $\sum^0 a_{jk}x_k$, where the non-singular constant $(n-1)$ -matrix (a_{jk}) depends on m_1, \dots, m_n in such a way that the momenta canonically conjugate to the coordinates $\sum^0 a_{jk}x_k$ become constant multiples of the respective velocities $\sum^0 a_{jk}x'_k$, while (2₁) or (7₁) is transformed into a diagonal form.

§385. By the barycentric chain belonging to the barycentric inertial position vectors ξ_1, \dots, ξ_n of m_1, \dots, m_n will be meant the sequences of $n-1$ three-vectors

$$(11) \quad X_j = \xi_{j+1} - \sum_{k=1}^j m_k \xi_k / \sum_{k=1}^j m_k, \quad (j = 1, \dots, n-1);$$

so that X_j is the position vector of m_{j+1} with reference to the centre of mass of the j bodies m_1, \dots, m_j . If one introduces the abbreviations

$$(12_1) \quad \mu_j = \sum_{k=1}^j m_k, \quad (\mu_0 = 0); \quad (12_2) \quad M_j = m_{j+1}\mu_j/\mu_{j+1}, \quad (M_0 = 0),$$

the connection between the X_j and the heliocentric x_j is given by the reciprocal pair of linear substitutions*

* It is understood that x_{j+1} in (13₁) denotes 0 if $j = n-1$ (cf. (8), §341), and that the first term on the right of (13₂) is missing if $j = 1$ (cf. (12₂)), finally that the summation on the right of (13₂) is vacuous if $j = n-1$.

$$(13_1) \quad X_j = x_{j+1} - \mu_j^{-1} \sum_{k=1}^j m_k x_k;$$

$$(13_2) \quad x_j = m_j^{-1} M_{j-1} X_{j-1} - \sum_{k=j}^{n-2} \mu_k^{-1} M_k X_k - X_{n-1},$$

where $j = 1, \dots, n-1$. For, on the one hand, it is clear from (1₁) and (12₁) that (13₁) is equivalent to the definition (11). And, on the other hand, (13₁), when combined with (12₁)–(12₂), implies the recursion formula $x_{j+1} - x_j = X_j - m_j^{-1} M_{j-1} X_{j-1}$ which, when telescoped from $x_n - x_{n-1} \equiv 0 - x_{n-1}$ onward, clearly leads to the inversion (13₂) of (13₁). The determinant of the linear substitution (13₁) is easily found to be $(-1)^{n-1}$.

Let (m_{jk}) denote the $(n-1)$ -matrix of the linear substitution (13₂); so that $x_j = \sum^0 m_{jk} X_k$, by (3₁). In view of (13₂) and (12₁)–(12₂), the coefficients m_{jk} are functions of the masses m_1, \dots, m_n alone and, as easily verified, satisfy the identities

$$(14) \quad \sum^0 m_i m_{lj} m_{lk} - (\sum^0 m_l m_{lj})(\sum^0 m_l m_{lk})/\mu = M_j e_{jk}, \quad (\sum^0 = \sum_{l=1}^{n-1}),$$

where (e_{jk}) is the unit matrix and $\mu = m_1 + \dots + m_n$. On substituting (13₂), i.e., $x_j = \sum^0 m_{jk} X_k$, into (4₁), (2₁), (4₂) and then using (14) in all three cases, one clearly obtains

$$(15_1) \quad J = \sum^0 M_j X_j^2; \quad (15_2) \quad T = \frac{1}{2} \sum^0 M_j X_j'^2;$$

$$(15_3) \quad \sum^0 M_j X_j \times X_j' = C.$$

§386. It follows that the barycentric chain (13₁) represents $n-1$ linear combinations of the $n-1$ heliocentric coordinate vectors x_j in such a way as to satisfy the requirement formulated at the end of §384.

In fact, the substitution (13₂), where the coefficients are the mass constants defined by (12₁)–(12₂), transforms the heliocentric Lagrangian function $L = T + U$ defined by (2₁)–(2₂), (1₂) into the Lagrangian function $L = T(X') + U(X)$ defined by the diagonal form (15₂) and the function U which one obtains by substituting (13₂) into (1₂). It follows, therefore, from §95 that the Lagrangian equations in terms of the X_j are $[L]_{X_j} = 0$, or, according to (15₂), simply

$$(16) \quad M_j X_j'' = U_{X_j}, \text{ where } U = U(X), \text{ by } (2_2), (1_2).$$

But, on placing $M_j X'_j = Y_j$, one has, from (15₂)–(15₃),

$$(17_1) \quad T = \frac{1}{2} \sum^0 M_j^{-1} Y_j^2; \quad (17_2) \quad \sum^0 X_j \times Y_j = C;$$

$$(17_3) \quad Y_j = M_j X'_j;$$

and (16) appears in the canonical form

$$(18) \quad \begin{aligned} Y'_j &= -H_{X_j}, \quad X'_j = H_{Y_j}, \quad \text{where} \\ H &= T - U \equiv \frac{1}{2} \sum^0 M_j^{-1} Y_j^2 - U(X). \end{aligned}$$

Notice that (18), (17₃) are of the same form as (9₁), (9₂), save that the n masses m_i are replaced by $n - 1$ masses M_j defined by (12₁)–(12₂), and the n barycentric inertial ξ_i by the $n - 1$ chained barycentric X_j of (11); while the barycentric invariant system (10₁)–(10₂) of (9₁) is eliminated, the degree of freedom of (18) being the same, $3(n - 1)$, as that of (8).

§387. Let, in particular, $n = 3$. Then (12₁)–(12₂), (13₁) reduce to

$$(19_1) \quad M_1 = \frac{m_1 m_2}{m_1 + m_2}, \quad M_2 = \frac{(m_1 + m_2) m_3}{m_1 + m_2 + m_3};$$

$$(19_2) \quad X_1 = x_2 - x_1, \quad X_2 = -\frac{m_1 x_1 + m_2 x_2}{m_1 + m_2}.$$

The inverse, (13₂), of the linear substitution (19₂) may be written as

$$(20_1) \quad x_j = (-1)^j \nu_j X_1 - X_2; \quad j = 1, 2;$$

$$(20_2) \quad \nu_1 : \nu_2 = m_2 : m_1; \quad \nu_1 + \nu_2 = 1.$$

Hence, (1₂) and (16) reduce to

$$(21_1) \quad \rho_{12} = |X_1|, \quad \rho_{j3} = |X_2 - (-1)^j \nu_j X_1|;$$

$$(21_2) \quad X''_j = p_{j1} X_1 + p_{j2} X_2,$$

if one uses (19₁) and the abbreviation

$$(22) \quad \begin{aligned} &\begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} \\ &= \begin{pmatrix} -(m_1 + m_2)/\rho_{12}^3 - m_3 \sum^0 \nu_j/\rho_{j3}^3 & m_3 \sum^0 (-1)^j/\rho_{j3}^3 \\ M_1 M_2^{-1} m_3 \sum^0 (-1)^j/\rho_{j3}^3 & -M_2^{-1} m_3 \sum^0 m_j/\rho_{j3}^3 \end{pmatrix}, \end{aligned}$$

where $\sum^0 \alpha_j = \alpha_1 + \alpha_2$, by (3₁). Also, from (15₃) and (15₁),

$$(23_1) \quad M_1(X_1 \times X'_1) + M_2(X_2 \times X'_2) = C;$$

$$(23_2) \quad J = M_1 X_1^2 + M_2 X_2^2,$$

while, by (2₂) and the relation (7₁) of §314,

$$(24_1) \quad U = m_1 m_2 / \rho_{12} + m_3 (m_1 / \rho_{13} + m_2 / \rho_{23}); \quad (24_2) \quad \frac{1}{2} J'' = 2h + U.$$

Needless to say, (21₂)–(22) are, in virtue of (19₁)–(21₁), identical with (15₂)–(16), §343 bis.

On applying (10₁) and (11) to the present case, $n = 3$, and then using (20₂), one sees that if a solution $X_1 = X_1(t)$, $X_2 = X_2(t)$ of (21₂) is known, the corresponding solution of the problem of $n = 3$ bodies in terms of the barycentric inertial positions $\xi_i(t)$ is given by

$$(25) \quad \begin{aligned} \xi_1 &= -\nu_1 X_1 - \mu^{-1} m_3 X_2, & \xi_2 &= \nu_2 X_1 - \mu^{-1} m_3 X_2, \\ \xi_3 &= (1 - \mu^{-1} m_3) X_2; & (\mu &= \sum m_i). \end{aligned}$$

Not only the three paths $\xi = \xi_i(t)$ of the m_i but also the two paths $\xi = X_j(t)$ of the hypothetical bodies (19₁) will be thought of as loci in the space $\xi = (\xi^I, \xi^{II}, \xi^{III})$.

§388. In this sense, one can speak of the tangent plane of the path of M_j through the centre of mass at a given t , that is to say of the plane Π_j^t in the ξ -space which goes through $\xi = 0$ and touches the path of M_j at a given t , where $j = 1, 2$. Thus, Π_j^t is the plane in the ξ -space which has the equation $(X_j(t) \times X_j'(t)) \cdot \xi = 0$, it being understood that Π_j^t does not exist if $X_j(t) \times X_j'(t) = 0$.

On multiplying (23₁) at a given t by an arbitrary ξ , one sees that $C \cdot \xi$ is a linear combination of the two $(X_j(t) \times X_j'(t)) \cdot \xi$. Hence, the intersection of the two planes Π_j^t belonging to one and the same t lies within the invariable plane $C \cdot \xi = 0$, provided that this intersection exists. If it exists, the solution $\xi_i = \xi_i(t)$ is, in view of (25), certainly not planar in the sense of §324; so that, by §326, the invariable plane must exist. Hence, if the integration constants of a solution $\xi_i = \xi_i(t)$ of the problem of $n = 3$ bodies are such that, except perhaps for isolated values of t ,

(I) both planes Π_j^t exist and (II) they are not parallel to each other, then $C \neq 0$, and the two planes Π_j^t intersect along a line N^t which rotates,* about the centre of mass, within the invariable plane.

* By this is not implied that N^t must actually depend on t . Incidentally, it seems to be quite a difficult problem, to determine all those particular solutions of the problem of $n = 3$ bodies which satisfy (I), (II) and have a line N^t of fixed position. As far as present knowledge goes, it is possible that no such solution exists. The question is connected with the one raised in §436 below.

§388 bis. There remain to be enumerated those particular solutions $\xi_i = \xi_i(t)$ of the problem of $n = 3$ bodies for which the conditions (I)–(II), §388 for the existence of the line N' are not satisfied. It is clear from (25) that condition (I) fails to be satisfied in case of the isosceles solutions enumerated in §346, as well as in the case of a rectilinear solution (§327). Similarly, it is clear from (25) that condition (II) fails to be satisfied by any planar solution (§324) and also by the non-planar isosceles solutions (i)–(ii), §346. It is not known whether or not (II) fails to be satisfied in case of some particular non-planar solutions distinct from these isosceles solutions.

Thus, the line N' does not exist in case of an arbitrary planar solution and in either case (i)–(ii), §346 of a non-planar isosceles solution. But it is an open question whether or not this enumeration is complete.

§389. The form (21₂) of the equations of the general problem of $n = 3$ bodies is well adapted to the treatment of isosceles solutions. In §345–§347, these solutions were treated only on the assumption* $m_1 = m_2$ of §344. It was mentioned at the end of §344 that this assumption is, as a matter of fact, a consequence of the definition† (§344) of an isosceles solution. For this fact, there is, in the main, only one proof known. And the details of this proof are too lengthy to be reproduced here. On the other hand, the underlying idea of the proof is simple enough. In order to indicate it, a few preparatory relations will be needed.

Suppose that, for arbitrarily given masses m_i , a given solution $\xi_i = \xi_i(t)$ of the problem of $n = 3$ bodies is such that $\rho_{13} = \rho_{23}$ for every t . Then, on squaring the relations (21₁), one readily finds that

$$(26_1) \quad 2X_1 \cdot X_2 = (\nu_2 - \nu_1)X_1^2;$$

$$(26_2) \quad \rho_{12}^2 = X_1^2; \quad (26_3) \quad \rho_{j3}^2 = X_2^2 + \nu_1 \nu_2 X_1^2,$$

where $j = 1, 2$. Furthermore, from (22) and (20₂), (19₁),

$$(27_1) \quad p_{11} = - (m_1 + m_2)/\rho_{12}^3 - m_3/\rho_{13}^3, \quad p_{22} = - \mu/\rho_{13}^3;$$

* On this assumption, (19₂) reduces to $X_1 = x_2 - x_1$, $X_2 = -\frac{1}{2}(x_1 + x_2)$ and is, save for the notation, identical with the substitution (20₁), §345 on which the treatment of the case $m_1 = m_2$ was based.

† Notice that this definition excludes the case of an equilateral solution, in which case the three m_i may be arbitrary, by §359 and §377 bis. Also notice that a collinear homographic solution satisfies the condition $\rho_{13} = \rho_{23}$ only when $m_1 = m_2$; cf. (12), §358.

$$(27_2) \quad p_{12} \equiv 0 \equiv p_{21},$$

where $\mu = m_1 + m_2 + m_3$. Put $|X_j| = r_j$, where $j = 1, 2$. Then

$$X_j^2 = r_j^2, \quad X_j \cdot X_j' = r_j r_j', \quad X_j'^2 + X_j \cdot X_j'' = r_j'^2 + r_j r_j'';$$

and

$$X_j \cdot X_j'' = p_{jj} r_j^2,$$

since $X_j'' = p_{jj} X_j$, by (21₂), (27₂). Thus, from (2), §65, where $a = X_j$, $b = X_j'$,

$$r_j^2(r_j r_j'' + r_j'^2 - p_{jj} r_j^2) = (r_j r_j')^2 + (X_j \times X_j')^2,$$

i.e., $r_j r_j'' = p_{jj} r_j^2 + (X_j \times X_j')^2 / r_j^2$. Since $X_j'' = p_{jj} X_j$ implies that $X_j \times X_j'' \equiv 0$, one has $X_j \times X_j' = A_j$, where $A_j = \text{const.}$ Thus,

$$(28_1) \quad \begin{aligned} r_1 r_1'' &= p_{11} r_1^2 + A_1^2 / r_1^2, \quad \text{where} \\ p_{11} &= - (m_1 + m_2) / r_1^3 - m_3 / (r_2^2 + \nu_1 \nu_2 r_1^2)^{\frac{1}{2}}, \end{aligned}$$

$$(28_2) \quad \begin{aligned} r_2 r_2'' &= p_{22} r_2^2 + A_2^2 / r_2^2, \quad \text{where} \\ p_{22} &= - (m_1 + m_2 + m_3) / (r_2^2 + \nu_1 \nu_2 r_1^2)^{\frac{1}{2}}; \end{aligned}$$

cf. (27₁) and (26₂)–(26₃), where $X_j^2 = r_j^2$. Finally, on substituting (26₂)–(26₃), where $X_j^2 = r_j^2$, into (23₂)–(24₁), one sees from (24₂) that

$$(29) \quad \frac{1}{2} (M_1 r_1^2 + M_2 r_2^2)'' = 2h + m_1 m_2 / r_1 + m_3 (m_1 + m_2) / (r_2^2 + \nu_1 \nu_2 r_1^2)^{\frac{1}{2}},$$

where M_j , ν_j are defined by (19₁), (20₂), and $h = \text{const.}$

Now, (28₁)–(28₂) is a pair of ordinary differential equations which, together with the 2 + 2 initial values $r_j(t^0)$, $r_j'(t^0)$, determine both functions $r_j(t)$ uniquely. But these functions must satisfy (29) also; so that the 2 functions $r_1(t)$, $r_2(t)$ are overdetermined by their 2 + 1 ordinary algebraic differential equations (28₁)–(29). Correspondingly, the underlying idea of the proof, mentioned at the beginning of the present article, may roughly be described as follows:

The analytic functions r_1 , r_2 of t are defined by (28₁)–(28₂) in such a way as to possess certain complex singularities. On the other hand, (29) also imposes on these functions certain complex singularities. These singularities, though not the functions $r_1(t)$, $r_2(t)$ themselves, may be determined a priori by explicit calculations, if recourse is made to the theory of analytic functions. And a detailed discussion shows that the singularities imposed on $r_1(t)$, $r_2(t)$ by (28₁)–

(28₂) are incompatible with those imposed by (29) unless either $r_1^2 \equiv r_2^2 + \nu_1 \nu_2 r_1^2$ or $m_1 = m_2$. Since in the first case the relations (26₂)–(26₃), where $X_j^2 = r_j^2$, show that $\rho_{12} \equiv \rho_{23} \equiv \rho_{31}$, it follows that $m_1 = m_2$ unless the solution is equilateral.*

It would, of course, be desirable to find a proof based on dynamical, rather than on function-theoretical, principles. But it is quite doubtful that such a proof exists. At any rate, the result is very deep, apparently much deeper than (ii), §371 (§374 bis notwithstanding).

§389 bis. There is a similar problem for any $n > 3$. In fact, if $n > 3$, all known flat but non-planar solutions of the problem of n bodies possess symmetries† similar to those possessed by the non-planar isosceles solutions (i)–(ii), §346. Hence, there arises the question of the *necessity* of these symmetry assumptions for the masses and the configurations in case of any flat but non-planar solution, if $n > 3$. This problem seems to be quite hard; it might depend on discussions of the type indicated in §389.

Elimination of the Angular Momentum

§390. Denoting by $\xi_i^\nu, \eta_i^\nu; C^\nu$, where $\nu = \text{I, II, III}$, the components of the 3-vectors $\xi_i, \eta_i; C$, and placing‡

$$(30) \quad F^\gamma = \sum (\xi_i^\alpha \eta_i^\beta - \xi_i^\beta \eta_i^\alpha),$$

where $(\alpha, \beta, \gamma) = (\text{I, II, III}), (\text{II, III, I}), (\text{III, I, II}),$

one can write the conservation relation (10₃) of the angular momentum of (9₁) in the form of the three scalar integrals $F^\nu = C^\nu$. Since it is readily verified from (30) that, in terms of the notation (19), §20, one has $(F^\beta; F^\alpha) = F^\gamma$, it follows from §92 that a Hamiltonian system cannot possess two of the three integrals $F^\nu = C^\nu$ without possessing the third also.§ Nevertheless, these three integrals are independ-

* Cf. the preceding footnote.

† Cf. the footnote to §325.

‡ It is of formal interest that the three components (30) of the angular momentum $\sum \xi_i \times \eta_i$ may be represented in terms of the $2n$ -matrix (16), §19 in the form of the three scalar products $F^\gamma = V^\beta \cdot IV^\alpha$, if V^ν denotes the $2n$ -vector which has the components $\eta_1^\nu, \dots, \eta_n^\nu; \xi_1^\nu, \dots, \xi_n^\nu$.

§ Actually, this situation is indicated by the proof which, in §316, established the three integrals $F^\nu = C^\nu$. In fact, if a system in a Euclidean 3-space is invariant under rotations about two of the coordinate axes, then it is in-

ent in the functional sense of §18, i.e., none of the three functions $F^\nu = F^\nu(\eta, \xi)$ is a function $f = f(F^\alpha, F^\beta)$ of the other two. In fact, partial differentiations show that the Jacobian of the three functions (30) with respect to the three variables $\eta_1^I, \eta_2^I, \eta_2^{II}$ (say) is not $\equiv 0$.

§391. Thus, on excluding from the $6n$ -dimensional (η, ξ) -space the lower-dimensional regions on which the three functions $F^\nu = F^\nu(\eta, \xi)$ become dependent in virtue of the vanishing of all 3-rowed Jacobians, and then attributing arbitrarily fixed values to the constant components C^ν of the angular momentum vector $(F^I, F^{II}, F^{III}) \equiv \sum \xi_i \times \eta_i = C$, one obtains a $(6n - 3)$ -dimensional region; so that the conservative system (9₁) of order $6n$ reduces to one of order $6n - 3$. Actually, it follows from the theory of Pfaffians that this conservative system of order $6n - 3$ is equivalent to one of order $6n - 4 = 2(3n - 2)$ and, what is more, to a conservative Hamiltonian system with $3n - 2$ degrees of freedom.

§391 bis. In order to obtain the latter system explicitly, the obvious procedure seems to be an adaptation of the idea applied above to the elimination of the centre of mass. In this regard, the sum of the n vectors $m_i \xi_i$, hence also the sum of the n vectors $m_i \xi_i' \equiv \eta_i$, is readily verified to vanish identically in virtue of (11), where the $n - 1$ vectors X_j are arbitrary. Thus, the reduction of the problem from the system (9₁) of order $6n$ to the system (18) of order $6(n - 1)$ is due to the fact that the 6 barycentric conditions (10₁)–(10₂), which represent an invariant system of (9₁), are parametrized in terms of the $6(n - 1)$ phase variables of (18) in such a way as to become identities. Correspondingly, the royal road leading to the reduction of §391 would be the introduction of suitable new phase variables in terms of which the invariant system (10₃) of (9₁), where $C = (C^I, C^{II}, C^{III})$ is fixed, becomes parametrized in such a way as to be satisfied identically. Of course, one would like this parametrization of (10₃) to be such that η_i^ν, ξ_i^ν appear as algebraic functions of symmetric structure in terms of the new variables. Unfortunately, no such algebraic parametrization of the quadratic 3-vector condition (10₃) has ever been devised, at least not for $n > 3$ (as to $n = 3$, cf. §394 below).

§392. Since the barycentric equations (9₁), (10₃), where $i = 1, \dots, n$,

variant under rotations about the third coordinate axis also, since every rotation about the origin may be composed of rotations about two perpendicular axes (cf. (21), §78).

are equivalent to the equations (18), (17₂), where $j = 1, \dots, n - 1$, all considerations of §390–§391 (and also the negative remark of §391 bis) remain valid if one replaces $\eta, \xi; n$ by $Y; X; n - 1$, respectively. In particular, the degree of freedom of the conservative Hamiltonian system mentioned at the end of §391 reduces to $3(n - 1) - 2 = 3n - 5$. If $n = 2$, this degree of freedom becomes 1 (which agrees, in view of §343, with (16₁)–(16₂), §214); while it becomes 4, if $n = 3$. Actually, the problem (9₁) of $n = 3$ bodies may be reduced by means of (10₁)–(10₃) to a conservative Hamiltonian system with 4 degrees of freedom. An explicit form of this reduced system belonging to $n = 3$ will be given below. If $n \geq 4$, no explicit representation of any merit seems to be known (cf. §391 bis).

§393. Consider first those solutions of the problem of $n = 3$ bodies which are collinear in the sense of §329. If such a solution is not rectilinear in the sense of §328, then it is, by §331, an homographic solution, and so §378 supplies the solution explicitly, making it dependent on the conservative dynamical system with a single degree of freedom which is defined by (21₁), §268. Hence, it is sufficient to consider the rectilinear case.

Then the barycentric inertial coordinate system $\xi = (\xi^I, \xi^{II}, \xi^{III})$ may be chosen so that $\xi_i^{II}(t) = 0 = \xi_i^{III}(t)$ for every i and t . Thus, (10₃) reduces to $0 = 0$, while the 3-vector ξ_i occurring in (9₁) may be considered to be a scalar ($= \xi_i^I$). Then (11) is a scalar for $j = 1, 2 (= n - 1)$, and so (18) becomes a conservative Hamiltonian system with $n - 1 = 2$ degrees of freedom.

Accordingly, the number 4 mentioned in §392 may be replaced by 2 in every collinear case, and by 1 in the collinear non-rectilinear case, of the problem of $n = 3$ bodies.

§394. Suppose, therefore, that the solution $\xi_i = \xi_i(t)$ under consideration is not collinear. Then it is clear from reasons of analyticity that syzygies (§327), if any, can occur only for isolated values of t and may, therefore, be disregarded. Thus, the $n = 3$ bodies m_i form a triangle $\Delta = \Delta(t)$. Let this triangle be oriented in such a way that the ordering m_1, m_2, m_3 of its vertices corresponds to the positive orientation, and let $\theta_i = \theta_i(t)$ denote the oriented exterior angle at the vertex m_i of $\Delta = \Delta(t)$. Then $\sum \theta_i = 0$ and, if $|\Delta| = |\Delta(t)|$ denotes the area of $\Delta = \Delta(t)$,

$$(31_1) \quad \sin \theta_i = \frac{2|\Delta|}{\rho_j \rho_k}, \quad \cos \theta_i = \frac{\rho_i^2 - \rho_j^2 - \rho_k^2}{2\rho_j \rho_k};$$

$$(31_2) \quad |\Delta| = \frac{\prod(\rho_j + \rho_k - \rho_i)^{\frac{1}{2}}}{4/(\sum \rho_i)^{\frac{1}{2}}} > 0,$$

where (i, j, k) runs through the three *cyclic* permutations of $(1, 2, 3)$ and ρ_i denotes the length of the side opposite the vertex m_i ; so that $\rho_i = \rho_{jk}$ in the notation (3₃), §314.

The barycentric position vectors of the vertices of $\Delta = \Delta(t)$ are the three $\xi_i = \xi_i(t)$; so that the plane $\Pi = \Pi(t)$ of the triangle, which always contains the point $\xi = 0$, varies with t , in general. If the solution $\xi_i = \xi_i(t)$ is such as to have no invariable plane, i.e., such that the angular momentum vector C vanishes, then the solution is, by §326, necessarily planar, and may, therefore, be assumed to take place within the (ξ^I, ξ^{II}) -plane. In this case, the oriented (ξ^I, ξ^{II}) -plane will be denoted by Π_* . If, on the other hand, $C \neq 0$, let Π_* denote the invariable plane $C \cdot \xi = 0$, which may be thought of as oriented by the normalization (7), §323 of the alternative sign in (6), §323. It is clear from (6), §323 that if the solution is planar, then Π_* coincides with the (ξ^I, ξ^{II}) -plane also when $C = 0$; so that $\Pi(t) = \Pi_*$ for every t . And Π_* is a well-defined plane of invariable inertial position through the centre of mass, whether the solution is planar or not.

Let $\iota = \iota(t)$ denote the inclination of the plane $\Pi = \Pi(t)$ of the triangle $\Delta = \Delta(t)$ of the three bodies towards this fixed plane Π_* . In particular, $\iota(t) = 0$ when all three m_i are in Π_* ; so that $\iota(t) \equiv 0$ if and only if the solution is planar.

The explicit form of the conservative Hamiltonian equations with $3n - 5 = 4$ degrees of freedom which were mentioned in §394 is now given in terms of the 4 coordinates $\iota; \rho_1, \rho_2, \rho_3$ by

$$(32) \quad I' = -H_\iota, \quad \iota' = H_I; \quad P_i' = -H_{\rho_i}, \quad \rho_i' = H_{P_i}, \quad (i = 1, 2, 3),$$

where $I; P_1, P_2, P_3$ denote the momenta canonically conjugate to these coordinates, and the Hamiltonian function is†

$$H \equiv H(I, P_1, P_2, P_3; \iota, \rho_1, \rho_2, \rho_3)$$

† Corresponding to the fact that the first term of (33) contains the factor $C^2 \sin^2 \iota$, this first term of H is thought of as being 0 when either $|C| = 0$ or $\sin \iota = 0$, although the term then contains the meaningless expression $\sin^2 (I/0 + \dots)$. But $|C| = 0$ implies that the solution is planar, i.e., that $\iota = 0$ for every t ; while if the solution is not planar, it is clear for reasons of analyticity that $\iota = \iota(t)$ cannot vanish except for isolated values of t , at most. Thus, the first term of H vanishes identically or becomes singular for isolated values of t , at most, according as the solution is or is not planar.

$$\begin{aligned}
(33) \quad &= \frac{|C|^2 \sin^2 \iota}{4|\Delta|} \sum \frac{\rho_i^2}{m_i} \sin^2 \left(\frac{I}{|C| \sin \iota} + \frac{\theta_j - \theta_k}{3} \right) \\
&+ \sum \frac{P_j^2 + P_k^2 - 2P_j P_k \cos \theta_i}{2m_i} \\
&+ |C| \cos \iota \sum \left(\frac{P_j}{\rho_k} - \frac{P_k}{\rho_j} \right) \frac{\sin \theta_i}{3m_i} \\
&+ |C|^2 \cos^2 \iota \sum \frac{\rho_j^2 + \rho_k^2 - \frac{1}{2}\rho_i^2}{36m_i \rho_j^2 \rho_k^2} - \sum \frac{m_j m_k}{\rho_i},
\end{aligned}$$

if Δ ; $\theta_1, \theta_2, \theta_3$ in (33) are thought of as expressed by means of (31₁)–(31₂) as functions of ρ_1, ρ_2, ρ_3 , if the value of the constant $|C| \geq 0$ is fixed, and if

$$(34_1) \quad \sum f_{ijk} = f_{123} + f_{231} + f_{312}; \quad (34_2) \quad \rho_i = |\xi_j - \xi_k| = \rho_{jk}.$$

The verification of the fact that the conservative Hamiltonian system (32)–(33) with 4 degrees of freedom is, in virtue of (10₁)–(10₃), equivalent to the conservative Hamiltonian system (9₁)–(9₂) with $3n = 9$ degrees of freedom requires only successive differentiations and substitutions. These elementary but lengthy calculations will be omitted. Incidentally, it turns out that the Hamiltonian functions H of (9₁) and of (32) are identical with each other in virtue of the geometrical (or, rather, kinematical) transformation formulae which connect the coordinates ι, ρ_i and the respective momenta I, P_i with the 3-vectors ξ_i and η_i .

§395. The Lagrangian function

$$(35) \quad L = L(\iota', \rho_1', \rho_2', \rho_3'; \iota, \rho_1, \rho_2, \rho_3)$$

belonging to (33) may be obtained from (2₁), §15, if one calculates the momenta $I; P_i$ in terms of the velocities ι', ρ_i' and coordinates ι, ρ_i (by applying (1₂), §15). However, the resulting representation of the momenta in terms of the velocities and the coordinates seems to be awkward and has never been used explicitly. At any rate, the 1 + 3 conservative Lagrangian equations $[L]_{\iota} = 0; [L]_{\rho_i} = 0$ represent the non-collinear problem of $n = 3$ bodies in terms of the inclination $\iota = \iota(t)$ of $\Pi = \Pi(t)$ towards the fixed plane Π_* (cf. §394), and of the 3 mutual distances $\rho_i = \rho_i(t)$ within the plane $\Pi = \Pi(t)$ of the 3 barycentric inertial vectors $\xi_i = \xi_i(t)$.

§396. It is quite interesting that only the single Eulerian angle ι occurs in the reduced problem. In fact, (23), §78 shows that in order to determine the relative position of the two planes $\Pi(t)$, Π_* , one needs, besides the inclination $\iota(t)$, also the node $\nu(t)$. Accordingly, a kinematical consequence of the possibility of a reduction of (9₁) to (32) is that a suitable application of the conservation of the angular momentum $\sum \xi_i \times \eta_i$ ($= C = \text{fixed constant}$) eliminates the node $\nu(t)$ of $\Pi(t)$ with reference to Π_* .

§397. Needless to say, the node ν also is needed in order to determine the three barycentric position vectors ξ_i . However, the elimination process which leads from (9₁)–(10₃) to (32) shows that, unless the solution is planar (in which case neither of the Eulerian angles ι , ν is needed), one has

$$(36) \quad \nu' = \frac{|C| \sum m_i}{|\Delta|} \sum \frac{\rho_i^2}{m_i} \sin^2 \left(\frac{1}{|C| \sin \iota} + \frac{\theta_j - \theta_k}{3} \right).$$

But if a solution of the reduced problem (32) is known, the expression on the right of (36) becomes in virtue of (31₁)–(31₂) a known function of t , and so $\nu = \nu(t)$ follows from (36) by a quadrature.

§398. The reduced degree of freedom, $3n - 5$, mentioned in §392, may be replaced by the smaller number $2n - 3$, if only planar solutions of the problem of n bodies are considered.

First, if the solution is collinear, an obvious repetition of the considerations of §393 shows that $3n - 5$ reduces to $n - 1$ or 1 according as the solution is or is not rectilinear. Suppose, therefore, that the planar solution is not collinear and choose its plane to be the (ξ^I, ξ^{II}) -plane. Then the ξ_i , and so, by (11), also the X_j , become 2-vectors. Hence, (18), where $j = 1, \dots, n - 1$, is a conservative dynamical system with $2n - 2$ degrees of freedom, and admits the integral $\sum^0 (X_j^I \dot{Y}_j^{II} - X_j^{II} \dot{Y}_j^I) = C^{III} = \pm |C|$ to which the three integrals represented by (17₂) reduce in virtue of $\xi_j^{III} \equiv 0 \equiv X_j^{III}$. Thus, elimination of the angular momentum C leads to a conservative system of order $2(2n - 2) - 1 = 4n - 5$. Actually, the Pfaffian implications mentioned in §391 show that this conservative system of order $4n - 5$ is equivalent to one of order $4n - 6 = 2(2n - 3)$ and, what is more, to a conservative Hamiltonian system with $2n - 3$ degrees of freedom.

§399. If $n = 3$, this system with $2n - 3 = 3$ degrees of freedom

follows from §394 in an explicit form. In fact, the footnote to §394 states that (33) reduces in the planar case to

$$(37) \quad H \equiv H(P_1, P_2, P_3; \rho_1, \rho_2, \rho_3) = \frac{1}{2} \sum m_i^{-1} (P_j^2 + P_k^2 - 2P_j P_k \cos \theta_i) \\ + |C| \sum \left(\frac{P_j}{\rho_k} - \frac{P_k}{\rho_j} \right) \frac{\sin \theta_i}{3m_i} - \sum \left(\frac{m_j m_k}{\rho_i} - \frac{\rho_j^2 + \rho_k^2 - \frac{1}{2}\rho_i^2}{36m_i \rho_j^2 \rho_k^2} |C|^2 \right),$$

where $\theta_i = \theta_i(\rho_1, \rho_2, \rho_3)$, by (31₁)–(31₂). Thus, the degree of freedom of the system (32) diminishes from 4 to 3. In fact, the first pair of the eight equations (32) reduces to $I = \text{const.}$, $\iota = \text{const.}$, the partial derivatives H_{ι} , H_I of (37) being $\equiv 0$.

Clearly, (37) appears in the form (7), §157, if one puts $P = p$, $\rho = q$ and

$$(38) \quad g^{ii} = m_j^{-1} + m_k^{-1}, \quad g^{ik} = -m_j^{-1} \cos \theta_j, \quad (i \neq j \neq k \neq i), \\ - \sum f^i(q) p_i - V(q) \text{ being represented by the last two terms} \\ (= |C| \sum \dots - \sum \dots) \text{ of (37); in particular, the condition} \\ (f^i) \equiv (0) \text{ of reversibility (§158) is satisfied only when } C = 0.$$

It is readily verified from (31₁)–(31₂) and (38) that the 3-matrix function $(g^{ik}) = (g^{ki})$ of (ρ_1, ρ_2, ρ_3) is everywhere positive definite, since $\rho_i < \rho_j + \rho_k$. In particular, the reciprocal matrix $(g_{ik}) = (g^{ik})^{-1}$ exists. Its elements are homogeneous functions of degree 0 in (ρ_1, ρ_2, ρ_3) , since the same holds, in view of (31₁) and (38), for the elements of (g^{ik}) .

§399 bis. Suppose, in particular, that $C = 0$ (so that the solution is necessarily planar). Then (37) simplifies to $H = T - U$, where $U = \sum m_j m_k / \rho_i$, while $T = \frac{1}{2} \sum \sum g^{ik} P_i P_k$, by (38). Hence, if the energy constant h is fixed, the problem is equivalent to the problem of geodesics on the 3-dimensional Riemannian manifold on which the square of the line element is given by (13), §179, where $q_i = \rho_i$. According to §178, the corresponding Lagrangian function and energy constant are, if $\tilde{g}_{ik} = 2(U + h)g_{ik}$,

$$(39) \quad \tilde{L} = \tilde{T} + \tilde{U} \text{ and } \tilde{h} = \frac{1}{2}, \text{ where } \tilde{T} = \frac{1}{2} \sum \sum \tilde{g}_{ik} \dot{\rho}_i \dot{\rho}_k, \tilde{U} \equiv 0,$$

the dots denoting differentiations with respect to the arc length along the geodesic.

§400. As an application, consider those non-collinear solutions of the problem of $n = 3$ bodies for which not only the angular mo-

mentum C but also the energy constant h vanishes. Then $\tilde{g}_{ik} = 2(U + h)g_{ik} = 2Ug_{ik}$, where $U = \sum m_j m_k / \rho_i$. Hence, the functions \tilde{g}_{ik} of (ρ_1, ρ_2, ρ_3) are homogeneous of degree $\alpha = -1$, since the g_{ik} are, by the end of §399, homogeneous of degree 0. Consequently, application of (19), §159 to (39), §399 bis gives

$$(\sum \sum \tilde{g}_{ik} \rho_i \dot{\rho}_k)' = (-1 + 2)(0 + \tfrac{1}{2}) + \sum 0 = \tfrac{1}{2};$$

i.e.,

$$\sum \sum \tilde{g}_{ik} \rho_i \dot{\rho}_k = \tfrac{1}{2}s + c,$$

where s is the arc length to which the dots refer, and c is an integration constant. Since $\tilde{g}_{ik} = 2Ug_{ik}$, it follows that

$$(40) \quad 2U \sum \sum g_{ik} \rho_i \dot{\rho}_k - \tfrac{1}{2}s = c, \quad (U = \sum m_j m_k / \rho_i).$$

Now, (40) is a non-conservative integral distinct from the energy integral $U \sum \sum g_{ik} \dot{\rho}_i \dot{\rho}_k = \text{const. } (= \tfrac{1}{2})$ of (39).

§401. The result (40), where $n = 3$, $C = 0$ and $h = 0$, has an analogue in the case where only $h = 0$ is assumed, while n and C are arbitrary (so that the solution need not be planar).

In fact, if the energy h is arbitrarily fixed (≥ 0), the problem $m_i \xi_i'' = U_{\xi_i}$ of n bodies is reduced by (13), §179 to the problem of geodesics on the $3n$ -dimensional Riemannian manifold on which the square of the line element is $\sum 2(U + h)m_i(d\xi_i)^2$, where ξ_i is a 3-vector and $U = \sum^* m_j m_k / |\xi_j - \xi_k|$. Hence, if $h = 0$, the coefficients of the $(d\xi_i)^2$ become homogeneous of degree $\alpha = -1$. Consequently, if $h = 0$, a repetition of the proof of (40) supplies the integral

$$(41) \quad 2U \sum m_i \xi_i \cdot \dot{\xi}_i - \tfrac{1}{2}s = c; \text{ while } U \sum m_i \dot{\xi}_i^2 \equiv \tfrac{1}{2}\dot{s}^2 = \tfrac{1}{2}, \quad (\dot{s} = 1).$$

§402. However, it would be a mistake to assume that this integral of the geodesic problem belonging to $h = 0$ contains anything new. In fact, (41) may be written as $U\dot{J} - \tfrac{1}{2}s = c$, $\dot{s} = 1$, where $J = \sum m_i \xi_i^2$. Hence, (41) is equivalent to $(U\dot{J})' = \tfrac{1}{2}$. But this relation is, in view of the connection between the time variables t and s ($= \bar{t}$; cf. (10), §176), identical with the relation to which (7₁), §315 reduces in the present case, $h = 0$.

It may be verified from (38) and (31₁) that also (40) is equivalent to $(U\dot{J})' = \tfrac{1}{2}$, since J may be represented by (12₂), §333, where $\rho_{jk} = \rho_i$.

§403. The conservation principles (10₁)–(10₃) imply for the general problem of $n = 3$ bodies an elementary geometrical fact which,

without any reference to (10₁)–(10₃), may be established as follows:

First, the three 3-vector equations $m_i \xi_i'' = U_{\xi_i}$ may be written in the form

$$(42_1) \quad \xi_i'' = \kappa \rho_i^3 (\xi_* - \xi_i);$$

$$(42_2) \quad \rho_i = |\xi_j - \xi_k|; \quad (i, j, k) = (1, 2, 3), (2, 3, 1), (3, 1, 2),$$

if the 3-vector $\xi_* = \xi_*(t)$ and the scalar $\kappa = \kappa(t)$ are defined by

$$(43_1) \quad \xi_* = \sum m_i \rho_i^3 \xi_i : \sum m_i \rho_i^3; \quad (43_2) \quad \kappa = \sum m_i \rho_i^3 : \prod \rho_i^3.$$

In fact, it is clear from (42₂) that the explicit form of $m_i \xi_i'' = U_{\xi_i}$, where $U = \sum m_j m_k / \rho_i$, is

$$(44) \quad \begin{aligned} \xi_i'' &= m_k \frac{\xi_k - \xi_i}{\rho_j^3} + m_j \frac{\xi_j - \xi_i}{\rho_k^3} \\ &\equiv \frac{\rho_k^3 m_k (\xi_k - \xi_i) + \rho_j^3 m_j (\xi_j - \xi_i)}{\rho_j^3 \rho_k^3}. \end{aligned}$$

Adding $\rho_i^3 m_i \sum (\xi_i - \xi_i) = 0$ to the last numerator and then using the abbreviations (43₁)–(43₂), one clearly obtains (42₁).

§404. According to (42₁), any of the $n = 3$ forces $U_{\xi_i} = m_i \xi_i''$ is, for every t and i , the product of a scalar function ($= -\kappa m_i \rho_i^3$) and of the 3-vector $\xi_i - \xi_*$, where $\xi_* = \xi_*(t)$ is independent of i . In other words, every solution $\xi_i = \xi_i(t)$ of the problem of $n = 3$ bodies is such that the $n = 3$ forces of gravitation which act on the three m_i are directed towards a certain point $\xi_* = \xi_*(t)$ of the ξ -space. This is the fact referred to at the beginning of §403.

The point $\xi_*(t)$ is, of course, uniquely determined except at dates $t = t^0$ of syzygies (§327), and is called the centre of force. Although (43₁) defines a unique point $\xi_*(t)$ also at dates $t = t^0$ of syzygies, the centre of force will not be considered as defined at such dates.

Needless to say, the centre of force has nothing to do with the centre of mass, which is $\xi = 0$ for every t . In fact, the centre of force is, by (43₁), the centre of mass not of the m_i but of three ideal masses $m_i \rho_i^3$ which have the same barycentric positions ξ_i as the m_i .

§405. Since dates of syzygies have been excluded, $\det(\xi_1, \xi_2, \xi_3) \neq 0$. Hence, if the three ξ_i are such that not only $\sum m_i \xi_i = 0$ but also $\sum m_i \rho_i^3 \xi_i = 0$, then $m_i \rho_i^3 : m_i$, i.e. ρ_i , is independent of i , and vice versa. This means, in view of (43₁) and (42₂), that the centre of

force is the centre of mass for those and only those t (if any) at which the triangle formed by the three masses happens to be equilateral.

An obvious extension of this consideration shows that one of three masses, say the mass at ξ_3 , is collinear with the centre of force and the centre of mass at those and only those dates t at which the condition for an isosceles triangle is satisfied; it being understood that (19₁), §344 need not hold.

§406. One can readily determine also those configurations for which the point (43₁) becomes the centre of mass in case of a syzygy (a case excluded in §404).

In fact, if the m_i are in syzygy, one can assume that the ξ_i are scalars ($= \xi_i^I$), and that $\xi_1 < \xi_2 < \xi_3$. Then $\xi_j - \xi_k = (-1)^i \rho_i$, by (42₂). Hence, by the last line of §322 bis,

$$\mu \xi_1 = -m_3 \rho_2 - m_2 \rho_3, \quad \mu \xi_2 = +m_1 \rho_3 - m_3 \rho_1, \quad \mu \xi_3 = m_2 \rho_1 + m_1 \rho_2;$$

$$(\mu = \sum m_i).$$

On substituting this representation of the ξ_i into (43₁) and noting that $\rho_2 = \rho_1 + \rho_3$, one readily finds that $\xi_* = 0$ if and only if the ratio $\lambda = \rho_3 : \rho_1$ is a root $\lambda = \lambda(m_1, m_2, m_3)$ of

$$(45) \quad m_1(m_2 + m_3)\lambda^4 + m_1(m_3 + 3m_2)\lambda^3 + 3m_2(m_1 - m_3)\lambda^2 \\ - m_3(m_1 + 3m_2)\lambda - m_3(m_1 + m_2) = 0$$

(and not, as one might have expected, of (11), §358).

Real Singularities

§407. Along a given solution of the problem of n bodies, put

$$(1) \quad r(t) \equiv r = \text{Min} (\rho_{12}, \rho_{13}, \dots, \rho_{n-1, n}), \quad (\rho_{jk} = \rho_{jk}(t)),$$

where $\rho_{jk} = |\xi_j - \xi_k|$ and, if $H = \frac{1}{2} \sum m_i^{-1} \eta_i^2 - U(\xi_1, \dots, \xi_n)$, then, by §320,

$$(2_1) \quad \eta'_i = -H_{\xi_i}, \quad \xi'_i = H_{\eta_i}; \quad (2_2) \quad \eta_i = \eta_i(t), \quad \xi_i = \xi_i(t).$$

By the elements of the theory of ordinary differential equations, every solution (2₂) of (2₁) depends analytically on t , since the partial derivatives of H (η_1, \dots, ξ_n) depend analytically on η_1, \dots, ξ_n .

Thus far it was always assumed that there is given on a certain, finite or infinite, t -interval a solution (2₂) of (2₁). We did not ask, on what t -interval does (2₂) exist, if it belongs to given initial values $\eta_i(t_0)$, $\xi_i(t_0)$ assigned to a fixed $t = t_0$. Nor did we ask, for what

finite t^* can at least one of the $6n$ analytic functions (2_2) acquire a singularity, and what is the function-theoretical character and dynamical significance of such a singular date t^* . In what follows, these fundamental questions will be considered.

Since the differential equations (2_1) are non-linear, the function-theoretical problem becomes hopelessly involved if one allows either unrestricted complex values of t or complex initial values assigned to a real t_0 . Hence, it will always be assumed that, on the one hand, there are assigned real initial values to a real t_0 , so that, H being real for real (η, ξ) , the solution (2_2) is real for real t ; and that, on the other hand, the functions (2_2) are thought of as continued analytically from t_0 onward along the real t -axis. This will, of course, involve certain complex t situated in a narrow domain about the real t -interval under consideration. Nevertheless, t will be understood to be real, unless the contrary is stated.

§408. The formal basis of the following considerations will be the remark that if μ denotes the total mass $\sum m_i$, and μ_0 the least among the $m_i > 0$, finally h the energy constant $H(\eta_1(t_0), \dots, \xi_n(t_0))$, then

$$(3) \quad |H_{\xi_i}(t)| < \{\mu/r(t)\}^2, \quad |H_{\eta_i}(t)| < \{(2/\mu_0^3)(|h| + \mu^2/r(t))\}^{\frac{1}{2}}$$

holds for every t on any t -interval on which none of the $\frac{1}{2}n(n-1)$ distances $\rho_{jk} = \rho_{jk}(t)$ vanishes. In fact, $H_{\xi_i} = -U_{\xi_i}$, where $U = \sum^* m_j m_k / \rho_{jk}$; so that the first inequality (3) is clear from (1). Similarly, $H_{\eta_i} = m_i^{-1} \eta_i$; so that the second inequality (3) follows from $\frac{1}{2} \sum m_i^{-1} \eta_i^2 = U + h$, since $0 < U < \mu^2/r$.

Let I denote the t -interval on which the solution (2_2) , defined by the initial conditions $\eta_i(t_0), \xi_i(t_0)$ which are assigned to an initial t_0 contained in I , is known to exist, to be regular analytic, and to be such that the minimum (1) of the $\frac{1}{2}n(n-1)$ mutual distances exceeds some fixed positive lower bound, r^* , for all t contained in I . Choose any fixed \bar{t} contained in I , and assign to (2_1) the initial conditions $\eta_i(\bar{t}), \xi_i(\bar{t})$. On applying the local existence and uniqueness theorem of regular analytic ordinary differential equations at the new initial date \bar{t} , one infers from (3) the existence of two positive numbers α^*, β^* which depend only on the given $r^* > 0$, the masses m_i , and the energy constant h , and have the following properties: As long as t is in the domain† $|t - \bar{t}| < \alpha^*$, the solution (2_2) exists,

† The facts stated, together with $(4_1)-(4_2)$, are true whether the domain $|t - \bar{t}| < \alpha^*$ is meant to be the real t -interval $\bar{t} - \alpha^* < t < \bar{t} + \alpha^*$ or the

is regular analytic, and such that

$$(4_1) \quad |\eta_i(t) - \eta_i(\bar{t})| < \beta^*, \quad |\xi_i(t) - \xi_i(\bar{t})| < \beta^*;$$

$$(4_2) \quad r(t) \geq \frac{1}{2}r^*; \text{ cf. (1).}$$

The point is that α^*, β^* do not depend on the choice of \bar{t} . Cf. also §79.

§409. Thus, it is clear (cf. §84) from the covering theorem of Heine-Borel that if the solution (2₂), defined by the initial conditions $\eta_i(t_0), \xi_i(t_0)$, either ceases to exist when t tends to some finite real $t = t^*$ or is such that at least one of the $6n$ analytic functions (2₂) has a singularity at a real $t = t^* \neq \infty$, then the positive function (1) of the real variable t must come arbitrarily close to 0 as $t \rightarrow t^*$. In other words, the lower limit $\underline{\lim} r(t) = 0$, when t tends increasingly or decreasingly to the critical t^* (according as $t_0 < t^*$ or $t_0 > t^*$). Since the time variable t may be replaced by $\pm t + \text{const.}$, one can assume without loss of generality that the initial $t_0 > 0$, and that the critical $t^* = 0$; so that $\underline{\lim} r(t) = 0$ as $t \rightarrow +0$.

Actually, not only $\underline{\lim} r(t) = 0$ but also $\lim r(t) = 0$. For suppose, if possible, that $\underline{\lim} r(t) = 0$ is compatible with $\overline{\lim} r(t) > 0$. Then there exist a number $r^* > 0$ and a sequence t_1, t_2, \dots such that $r(t_m) > r^*$ for every m , while $t_m \rightarrow +0$ as $m \rightarrow +\infty$. Hence, on placing $\bar{t} = t_m$ for an arbitrarily fixed m , and then applying the facts mentioned at the end of §408, one sees that (4₂) holds for every t contained in the interval $|t - t_m| < \alpha^*$, where $\alpha^* > 0$ is independent of m . Hence, if m is chosen so large that the t -interval $|t - t_m| < \alpha^*$ contains the point $\lim t_m = 0$, it follows that $r(t) \geq \frac{1}{2}r^*$ for every t sufficiently close to $t = 0$; so that $\underline{\lim} r(t) \geq \frac{1}{2}r^*$. Since this contradicts the assumptions that $\underline{\lim} r(t) = 0$ and $r^* > 0$, the proof of $\lim r(t) = 0$ is complete.

§410. It is easy to see that for at least one of the analytic functions $\xi_i(t)$ to acquire a singularity as $t \rightarrow +0$, it is not only necessary (§409) but also sufficient that $r(t) \rightarrow 0$. For if $r(t) \rightarrow 0$, then, since $U = \sum^* m_j m_k / \rho_{jk}$ and $\frac{1}{2} \sum m_i \xi_i'^2 - U = h = \text{Const.}$, one sees from (1) that $\lim |\xi_i'(t)| = \infty$ for at least one i . This implies that the

complex t -circle, of radius α^* , about the point \bar{t} of the real axis of the t -plane. In the latter case, $\rho_{jk}(t)$ in (1) must be replaced by the square root of the square sum of the absolute values of the three complex numbers $\xi_i^\nu(t) - \xi_k^\nu(t)$, where $\nu = \text{I, II, III}$ in the notations of §313.

corresponding $\xi_i(t)$ must become singular at $t = 0$ (although it can tend to a finite limit as $t \rightarrow 0$; cf. the function $t^{\frac{2}{3}}$).

§411. Unfortunately, nothing is known as to the function-theoretical character of these singularities, if $n > 3$. The trouble starts with the lack of an adequate kinematical interpretation of the necessary and sufficient condition $\lim r(t) = 0$.

In view of (1), one might be tempted to interpret this condition by saying that some of the n bodies collide when $t \rightarrow +0$. But the legitimacy of this interpretation can be proved to-day only for $n \leq 3$ (cf. §365–§367; if $n = 2$, the situation is obvious from §343, since $m_1\xi_1 + m_2\xi_2 = 0$). The difficulty is that (1) might tend, as $t \rightarrow +0$, to 0 also when none of the mutual distances tends to 0, since the rôle of being the least among the $\frac{1}{2}n(n-1)$ numbers $\rho_{jk}(t)$ might be exchanged between them infinitely often, as $t \rightarrow +0$ (cf. the non-Newtonian example of §374 bis). In other words, $r(t) \rightarrow 0$ implies a “collision” only if it is known that the mutual distances must tend to limits. And this is, to-day, undecided for every $n > 3$. Even if it were decided, it would still not follow that the “collision” must take place at a definite point of the barycentric inertial coordinate system ξ , since a proof for the existence of the limiting positions $\lim \xi_i(t)$ would still be missing (cf. §365).

§411 bis. It is not even known whether or not all $|\xi_i(t)| < \text{const.}$, if $r(t) \rightarrow 0$, as $t \rightarrow +0$. All that is obvious is that $J = \sum m_i \xi_i^2$ must tend to a limit (≥ 0) which might be $+\infty$. In fact, since $J'' = 2U + 4h$, where $U = \sum^* m_j m_k / \rho_{jk}$ and $h = \text{const.}$, it is clear from (1) that $r(t) \rightarrow 0$ is equivalent to $J''(t) \rightarrow +\infty$. Thus, the function $J = J(t)$ is ultimately convex and tends, therefore, to a limit $\leq +\infty$.

§412. From now on it will be assumed that $n = 3$; so that

$$(5_1) \quad r(t) \equiv r = \text{Min} (\rho_{12}, \rho_{13}, \rho_{23}); \quad (5_2) \quad \rho_{ik} \leq \rho_{ij} + \rho_{jk};$$

(5₂) being the inequality between the sides of the triangle (m_1, m_2, m_3) which may be a segment.

It will first be shown that if (5₁) satisfies the condition $r(t) \rightarrow 0$, $t \rightarrow +0$, of §409–§410, then at least one of the three distances $\rho_{jk}(t)$ tends to 0.

Suppose, if possible, that all three $\overline{\rho_{jk}(t)} > 0$. Then (5₁) can tend to 0 only if at least two of the three ρ_{jk} , say ρ_{13} and ρ_{23} , interchange, as $t \rightarrow +0$, infinitely often the rôle of being the least among

the three ρ_{jk} at a fixed t . Let t_1, t_2, \dots denote dates at which this rôle is interchanged between ρ_{13} and ρ_{23} ; so that $\rho_{13}(t)$ and $\rho_{23}(t)$ have the common value (5_1) for every $t = t_m$, while $t_m \rightarrow +0$ as $m \rightarrow \infty$. Since $r(t) \rightarrow 0$ as $t \rightarrow +0$, it follows that $\rho_{13}(t_m) = \rho_{23}(t_m)$ tends, as $m \rightarrow \infty$, to 0. This implies, by (5_2) , that also $\rho_{12}(t_m) \rightarrow 0$. Since all three $\rho_{jk}(t_m) \rightarrow 0$, one sees from (12_2) , §333 that $J(t_m) \rightarrow 0$ as $m \rightarrow \infty$. Hence $\lim J(t) = 0$ as $t \rightarrow +0$. Since, by §411 bis, one always has $\lim J(t) = \lim J(t)$, it follows that $J(t) \rightarrow 0$ as $t \rightarrow +0$. This means, by (12_2) , §333, that all three $\rho_{jk}(t) \rightarrow 0$ and contradicts, therefore, the assumption that all three $\lim \rho_{jk}(t) > 0$.

This proves that at least one $\rho_{jk}(t) \rightarrow 0$. Choose the notations so that

$$(6) \quad \rho_{12}(t) \rightarrow 0; \quad \text{so that} \quad \rho_{13}(t) - \rho_{23}(t) \rightarrow 0, \quad \text{by} \quad (5_2).$$

§412 bis. It follows that the limit $\lim J(t) \leq +\infty$, established at the end in §411 bis, cannot be $= +\infty$.

Suppose, if possible, that the limit of (12_2) , §333 is $+\infty$. Then (6) implies that $\rho_{13}(t) \rightarrow +\infty$, $\rho_{23}(t) \rightarrow +\infty$; so that m_3 does not come arbitrarily close to m_1 and m_2 as $t \rightarrow +0$. Since $\rho_{12}(t) \rightarrow 0$, it follows that m_1 and m_2 participate in a binary collision as defined in §349. Hence, §352 is applicable and shows that there exist finite limits for all three $\xi_i(t)$, hence also for $\rho_{13}(t) = |\xi_1(t) - \xi_3(t)|$.

This contradiction proves that $\lim J(t) < +\infty$. Consequently, on using (12_2) , §333 again, one sees from (6) that $\rho_{13}(t)$ and $\rho_{23}(t)$ tend to a common finite limit ≥ 0 .

§413. On comparing this result with §409, one sees that if a solution $\xi_i = \xi_i(t)$ of the problem of $n = 3$ bodies exists and is regular analytic for small $t > 0$, and if at least one component of at least one of the three 3-vectors $\xi_i(t)$ acquires a singularity at $t = 0$, then there are, as $t \rightarrow +0$, only two cases possible: Either all three $\rho_{jk}(t)$ tend to 0 or one of them tends to 0 while the other two tend to a common finite positive limit. In other words, one has either a simultaneous collision in the sense of §335 or a binary collision in the sense of §349.

In the first case, all three $\xi_i(t)$ tend to the centre of mass $\xi = 0$ in a way required by §367 and (24) , §366. In the second case, none of the ξ_i can tend to $\xi = 0$ and, if the notation is chosen so that the binary collision takes place between m_1 and m_2 , there exist, by §352, finite limits

$$(7_1) \quad 0 \neq \xi_1^0 = \xi_2^0 \neq \xi_3^0 \neq 0; \quad (7_2) \quad \xi_3'^0; \quad (7_3) \quad \rho_{13}^0 = \rho_{23}^0 (> 0),$$

where the superscript ⁰ refers to $\lim t = +0$; furthermore, by (29), §350 and (28₃), §349,

$$(8_1) \quad \rho_{12} \sim \mathfrak{u} t^{\frac{1}{2}}, \quad \text{where} \quad \mathfrak{u} = \left[\frac{g}{2}(m_1 + m_2) \right]^{\frac{1}{2}};$$

$$(8_2) \quad \frac{1}{2} X_1'^2 |X_1| \rightarrow m_1 + m_2; \quad (|X_1| = \rho_1),$$

if X_1 denotes the relative position vector $\xi_2 - \xi_1$.

In the first case, one has $C = 0$, by §335; so that the solution must be planar, by §326. In the second case, C may but need not vanish, and, if $C \neq 0$, the solution may but need not be planar; finally, if it is (as it is in general) non-planar, all three m_i tend, by §353, to positions situated within the invariable plane.

§414. There arises the question whether or not the three analytic 3-vector functions $\xi_i(t)$, where $0 < t \rightarrow +0$, admit of a real analytic continuation through the date $t = 0$ of collision for small negative t ; cf. §268–§269. It will be shown that the answer to this question is always affirmative in case of a binary collision (§415–§420), while it may but need not be affirmative in case of a simultaneous collision (§421–§424); in which case the answer depends on the numerical values of the masses m_i and of the integration constants.

In §268–§269, the local uniformizing variable was the eccentric anomaly u , which is, by (3₂), §259, proportional to the undetermined integral of the reciprocal distance. Hence, the heuristic remarks of §349 suggest that in case of a binary collision one should try to regularize the singularity by introducing, instead of t , the independent variable

$$(9) \quad u \equiv u(t) = \int_0^t d\tilde{t} / \rho_{12}(\tilde{t}), \quad (\rho_{12} = |\xi_1 - \xi_2|).$$

According to (8₁), the integrand of (9) becomes infinite in the integrable order $\frac{2}{3}$, as $t \rightarrow +0$; so that (9) exists for $t > 0$, and is such that

$$(10) \quad u \sim 3\mathfrak{u}^{-1} t^{\frac{1}{2}} \quad \text{as} \quad t \rightarrow 0; \quad (\mathfrak{u} > 0).$$

§414 bis. In case of a simultaneous collision, not only one but each of the three reciprocal distances becomes infinite in the integrable order $\frac{2}{3}$; cf. §364 and (24), §365 bis. Since (7₃)–(8₁) are valid in case of a binary collision, it follows that, whether the collision is simultaneous or binary, $U(t) \equiv U = \sum^* m_j m_k / \rho_{jk}$ becomes infinite in the integrable order $\frac{2}{3}$; so that the same holds, in view of the identity

$J'' = 2U + 4h$, for $J'' = J''(t)$ also. Consequently, whether the collision is simultaneous or binary, $J' = J'(t)$ tends to a finite limit.

The Function-Theoretical Character of the Collisions

§415. Leaving aside, for a moment, the problem of singularities, collect the formulae belonging to (18), §386, if $n = 3$. Thus,

$$(11_1) \quad Y'_j = -H_{X_j}, \quad X'_j = H_{Y_j};$$

$$(11_2) \quad H = \frac{1}{2}M_1^{-1}Y_1^2 + \frac{1}{2}M_2^{-1}Y_2^2 - \sum^* m_l m_k / \rho_{lk},$$

where $j = 1, 2$ ($= n - 1$) and, by (20₂)–(21₁), §387,

$$(12_1) \quad \nu_j = (\mu_2 - m_j)/\mu_2, \quad \mu_2 = m_1 + m_2;$$

$$(12_2) \quad \rho_{j3} = |X_2 - (-1)^j \nu_j X_1|; \quad (12_3) \quad \rho_{12} = |X_1|.$$

Finally, (19₁), §387 and (25), §387 may be written as

$$(13_1) \quad M_1 = \nu_j m_j, \quad M_2 = \mu_2 m_3 / \mu;$$

$$(13_2) \quad \xi_j = (-1)^j \nu_j X_1 - \mu^{-1} m_3 X_2; \quad (13_3) \quad \xi_3 = (1 - m_3/\mu) X_2.$$

Next, introduce in place of the four 3-vectors Y_j, X_j four 3-vectors P_j, Q_j by

$$(14_1) \quad P_1 = Y_1/Y_1^2, \quad Q_1 = Y_1^2 X_1 - 2(Y_1 \cdot X_1) Y_1;$$

$$(14_2) \quad P_2 = Y_2, \quad Q_2 = X_2.$$

Since (14₁) is, save for the notation, the completely canonical, involutory transformation of §50, while (14₂) is the identical transformation, the transformation (14₁)–(14₂) is, by §33, completely canonical and involutory. Thus, (14₁)–(14₂) has the inverse

$$(15_1) \quad Y_1 = P_1/P_1^2, \quad X_1 = P_1^2 Q_1 - 2(P_1 \cdot Q_1) P_1;$$

$$(15_2) \quad Y_2 = P_2, \quad X_2 = Q_2,$$

and transforms (11₁)–(11₂) into

$$(16_1) \quad P'_j = -H_{Q_j}, \quad Q'_j = H_{P_j}; \quad (16_2) \quad H = H(P_1, P_2, Q_1, Q_2).$$

In order to compute (16₂), notice that, by (15₁)–(15₂),

$$(17) \quad \frac{1}{4}\kappa = \det \begin{pmatrix} \frac{1}{2}P_1 \cdot P_1 & P_1 \cdot Q_1 \\ P_2 \cdot Q_2 & Q_1 \cdot Q_2 \end{pmatrix}, \quad \text{if } \kappa = 2X_1 \cdot X_2;$$

and that (as already mentioned at the end of §50) one has

$$(18_1) \quad P_1^2 Y_1^2 = 1; \quad (18_2) \quad P_1^2 Q_1^2 = Y_1^2 X_1^2; \quad (18_3) \quad P_1 \cdot Q_1 + Y_1 \cdot X_1 = 0$$

in virtue of (14₁). From (18₁)–(18₂); (14₂); (12₂),

$$(19_1) \quad X_1^2 = (P_1^2)^2 Q_1^2; \quad (19_2) \quad X_2^2 = Q_2^2;$$

$$(19_3) \quad \rho_{j3}^2 = X_2^2 - (-1)^j \nu_{j\kappa} + \nu_j^2 X_1^2.$$

On substituting the two Y^2 from (18₁), (15₂) and the three $\rho = \rho(P, Q)$ from (12₃), (19₁)–(19₃) into (11₂), one obtains for (16₂) the explicit representation

$$(20) \quad H = \frac{1/P_1^2}{2M_1} + \frac{P_2^2}{2M_2} - \frac{m_1 m_2}{P_1^2 |Q_1|} - \sum_{j=1}^2 \frac{m_3 m_j}{\{Q_2^2 - (-1)^j \nu_{j\kappa} + (\nu_j P_1^2 |Q_1|)^2\}^{\frac{1}{2}}}.$$

§416. Clearly, (15₁)–(15₂) is an adaptation to the present case of the canonical extension of the coordinate transformation (24), §54, used in §259.

In order to make the analogy with §259 complete, consider those solutions of (11₁) which belong to an arbitrarily fixed value of the energy constant h , and then introduce instead of t the new time variable (9). On denoting by dots total differentiations with respect to this $u = u(t)$, one sees by applying the rule of §180 to $\bar{t} = u$, $G = \rho_{12}$ that, along every solution of energy h , the relations (16₁)–(16₂) can be replaced by

$$(21_1) \quad \dot{P}_j = -\bar{H}_{Q_j}, \quad \dot{Q}_j = \bar{H}_{P_j}; \quad (21_2) \quad \bar{H} = (-h + H)\rho_{12}.$$

Denoting by P_j^λ , Q_j^λ , where $\lambda = \text{I, II, III}$ and $j = 1, 2$, the components of the four 3-vectors P_j , Q_j , and expressing ρ_{12} by means of (12₃) and (19₁) in terms of Q_1 , P_1 , one can write (21₂) as

$$(22_1) \quad \bar{H} = \bar{H}(P_1^{\text{I}}, \dots, P_2^{\text{III}}, Q_1^{\text{I}}, \dots, Q_2^{\text{III}}); \quad (22_2) \quad \rho_{12} = P_1^2 |Q_1|,$$

the energy constant h having a fixed value.

According to (22₂), (21₂), (20), the explicit form of the function (22₁) of twelve scalar variables is

$$(23) \quad \begin{aligned} \bar{H} = |Q_1| & \left(-hP_1^2 + \frac{1}{2M_1} + \frac{P_1^2 P_2^2}{2M_2} \right. \\ & \left. - m_3 P_1^2 \sum_{j=1}^2 \frac{m_j}{\{Q_2^2 - (-1)^j \nu_j \kappa + (\nu_j P_1^2 |Q_1|)^2\}^{\frac{1}{2}}} \right) - m_1 m_2, \end{aligned}$$

where $\frac{1}{4}\kappa$ is an abbreviation for the determinant (17), the scalars ν_j , M_j defined by (12₁), (13₁) depend only on the fixed masses m_i , and finally

$$(24_1) \quad Q_1^2 = (Q_1^I)^2 + (Q_1^{II})^2 + (Q_1^{III})^2;$$

$$(24_2) \quad Q_2^2 = (Q_2^I)^2 + (Q_2^{II})^2 + (Q_2^{III})^2; \dots$$

§417. The isoenergetic canonical system (21₁)–(22₁), which is valid along any solution $\xi_i = \xi_i(t)$ of given energy h , will now be applied to a binary collision of m_1 and m_2 . Thus, if this collision takes place when t tends decreasingly to 0, one has $\rho_{12} \rightarrow 0$ as $t \rightarrow +0$, while ρ_{13} , ρ_{23} tend to a common positive limit, say α .

Using, instead of t , the time variable (9) of (21), one has $u \rightarrow +0$, instead of $t \rightarrow +0$. The given solution of (21₁) determines for every $u > 0$ a point

$$(25) \quad (P_1^I, \dots, P_2^{III}, Q_1^I, \dots, Q_2^{III})$$

in the twelve-dimensional phase-space. It will be shown that, as $u \rightarrow +0$, the point (25) remains in a closed bounded region which is entirely within the domain of regular analyticity of the analytic (but not everywhere regular) function (22₁) of twelve independent variables P_1^I, \dots, Q_2^{III} .

§417 bis. In order to prove this, it will be sufficient to show that, as $u \rightarrow +0$, both P_j and both Q_j remain bounded, and one has

$$(26_1) \quad P_1^2 |Q_1| \rightarrow 0;$$

$$(26_2) \quad \kappa \rightarrow 0;$$

$$(26_3) \quad |Q_2| \rightarrow \alpha > 0;$$

$$(26_4) \quad |Q_1| \rightarrow \beta > 0$$

for suitable α, β . For then it obviously follows from (23) and (24₁)–(24₂) that, as $u \rightarrow +0$, the point (22₁) does not come close to a singular point of the function (22₁) of twelve independent variables; κ being, by (17), a polynomial in these variables.

First, it will be shown that both pairs P_j, Q_j remain bounded as $u \rightarrow +0$. Since also (26₁)–(26₄) have to be proved, and since (26₄), (26₃) and (26₁), (26₄) imply the boundedness of Q_1, Q_2 and P_1 , respectively, it is sufficient to consider P_2 . But P_2 is, by (14₂) and (11₁)–

(11₂), identical with $Y_2 = M_2 X'_2$, where M_2 is a positive constant, while $X'_2 = \text{const. } \xi'_3$, by (13₃); and ξ'_3 remains bounded, since it tends to the finite limit (7₂). Thus, only (26₁)–(26₄) remain to be proved.

Next, (22₂) shows that (26₁) is true, by the assumption $\rho_{12} \rightarrow 0$. Furthermore, $\kappa = \rho_{13}^2 - \rho_{23}^2 + (\nu_2^2 - \nu_1^2)\rho_{12}^2$, by (19₃), (12₁) and (12₃); so that (26₂) follows from the fact that ρ_{13} and ρ_{23} tend to a common limit α , while $\rho_{12} \rightarrow 0$ and $\nu_j = \text{const.}$ Since α is positive by assumption, one sees from (19₁)–(19₃) that (26₃) is implied by (26₁)–(26₂). Finally, $X'_1 = M_1^{-1}Y_1$ and $Y_1^2|X_1| = |Q_1|$, by (11₁), (11₂) and (19₁), (18₁), respectively; so that (26₄) is, in view* of (8₂), satisfied by $\beta = 2(m_1 + m_2)M_1^2$.

§418. This completes the proof of the fact announced at the end of §417. But the derivatives of an analytic function can become singular only at the singular points of this function. Hence, if f_1, \dots, f_{12} are the first partial derivatives of (22₁), and **D** denotes a closed bounded region in the twelve-dimensional phase space $(P_1^I, \dots, Q_2^{III})$, it follows that the point (25) which represents the given solution of (21₁) at a fixed $u > 0$ remains, as $u \rightarrow +0$, entirely within a suitably chosen region **D** which is such as to contain none of the singularities of the twelve analytic functions f_1, \dots, f_{12} of twelve independent variables. But these f constitute, up to sign, the right-hand members of the equations (21₁). Hence, on combining the covering theorem of Heine-Borel with the local existence and uniqueness theorem of ordinary regular differential equations, one readily sees that any of the twelve scalar functions (25) of u , which represent the given solution of (21₁), must tend to a finite limit as $u \rightarrow +0$. Finally, this limiting position of the point (25) is again in the closed bounded region **D**. Consequently, application of the local existence and uniqueness theorem at $u = 0$ shows that all twelve functions (25) of u remain regular at $u = 0$.

Thus, the four 3-vector functions $P_j(u)$, $Q_j(u)$, where $j = 1, 2$, may be developed at $u = 0$ into power series which converge for sufficiently small $|u|$, represent the given solution of (21₁) for $u > 0$, and have, of course, real coefficients.

§419. Substitute these expansions of P_j , Q_j into the representation (15_j) of X_j , and then the resulting expansions of X_1 , X_2 into the

* The definition of X_1 after (8₂) was $X_1 = \xi_2 - \xi_1$. This agrees with (13₂), since $\nu_1 + \nu_2 = 1$, by (12₁).

representations (13₂)–(13₃) of the three ξ_i . Since these operations require only a finite number of additions and multiplications of power series, it follows that all three barycentric inertial position vectors ξ_i may be developed according to powers of u into regular power series with real coefficients. But the assumption was that there is, as $t \rightarrow +0$, a binary collision. Hence, on substituting $|\xi_1 - \xi_2| = \rho_{12} \equiv \rho_{12}(u)$ into (8₁), where $u > 0$, one sees from the definition (9), where $t > 0$, $u > 0$, that $t = t(u)$ may be developed at $u = 0$ into a regular power series which has real coefficients, vanishes at $u = 0$ in the third order (i.e., so that $t(u) = u^3 p(u)$, where $p(0) \neq 0$), and represents for small $u > 0$ the unique real inverse of the function (9), originally given for small $t > 0$; ($u > 0$).

Now define the three $\xi_i = \xi_i(u)$ and $t = t(u)$ for small $u < 0$ by their power series. Since the coefficients of these power series are real, and since $t(u)$ vanishes at $u = 0$ in the third order, it follows that the $\xi_i = \xi_i(t)$ are then uniquely defined for small $t < 0$ as real analytic continuations of the functions $\xi_i = \xi_i(t)$ which were originally given for small $t > 0$. In fact, $t(u) = u^3 p(u)$, where $p(0) \neq 0$, implies that the local inversion $u = u(t)$ of $t = t(u)$ may be developed into a real power series in $\sqrt[3]{t}$; so that one can define the $\xi_i(t)$ in terms of the $\xi_i(u)$ by placing $\xi_i(t) = \xi_i(u(t))$ for every t of small absolute value. Then, for reasons of analyticity, (2₁), where $\eta_i = m_i \xi_i'$, is satisfied for $t < 0$ also.

§420. Thus, the singularities mentioned in §410 are, in case of a binary collision at $t = 0$, algebraic singularities, with $\sqrt[3]{t}$ as local uniformizing variable; so that the situation is the same as in the elementary case analyzed in §268–§269.

Incidentally, a straightforward perusal of the above proof shows that $\xi_3(t)$ remains regular at $t = 0$, if m_3 is the body which does not participate in the collision.

§420 bis. If one wants a local uniformizing variable u valid for any pair of the colliding bodies and for any date of collision, (9) can be replaced by

$$(26) \quad u = \int^t U(\tilde{t}) d\tilde{t} \equiv \int^t (\sum^* m_j m_k / \rho_{jk}) d\tilde{t},$$

say. In fact, two of the three distances tend, in case of a binary collision at $t = 0$, to a positive limit; so that one has again $t(u) = u^3 p(u)$, $p(0) \neq 0$.

Notice that these choices of the time variable u in the present problem are equivalent to the choice of the time variable \bar{t} in the problem of §203–§205, where $dt/d\bar{t}$ is proportional to the product $r_1 r_2$.

§421. The proof of the statement of §414 concerning binary collisions is now complete. For the remaining case of simultaneous collisions, the statement of §414 was two-fold; namely, that in this case the situation

(i): can be, but (ii): is not always, the same as in §269 or §420.

The proof of (i) is supplied, for arbitrarily given values of the $n = 3$ masses m_i , by the example of those homographic (collinear or equilateral) solutions $\xi_i = \xi_i(t)$ which do not have an invariable plane ($C = 0$). In fact, §378 shows that these solutions always exist and present precisely the elementary problem treated in §268–§269.

The proof of (ii) will, in §422–§424, be supplied by showing that there exist, for arbitrarily given values of the m_i , solutions of the form

$$(27) \quad \xi_i(t) = t^{\frac{3}{2}} \sum_{n=0}^{\infty} \alpha_{in} t^{-sn}, \quad (0 < t < \text{const.}),$$

where the 3-vectors $\alpha_{i0}, \alpha_{i1}, \alpha_{i2}, \dots$ are, for all three values of i , real coefficients depending on the integration constants and not all three α_{i1} vanish; that the power series $t^{-\frac{3}{2}} \xi_i(t)$ in t^{-s} have non-vanishing radii of convergence; finally, that s is a negative number which depends only on the given values of the three m_i and is, as a matter of fact, an algebraic function of the m_i and is not independent of the m_i . Thus, the number $s < 0$ is rational only for exceptional values of the given m_i . That the existence of these solutions (27) will imply the proof of (ii), is seen as follows:

Clearly, the inertial barycentric position vectors (27) tend with $t(> 0)$ to 0; so that there is a simultaneous collision of all $n = 3$ bodies, as $t \rightarrow +0$. Choose the three masses m_i so that the positive number $s = s(m_1, m_2, m_3)$ is irrational. Then, since not all three α_{i1} vanish, at least one of the analytic 3-vector functions (27) of t has at $t = 0$ an isolated but essential (logarithmical) singularity. Thus, while the solution (27) is real for $t > 0$ and possesses infinitely many direct analytic continuations for $t < 0$, each of the resulting branches turns out to be complex for $t < 0$ (cf. §269–§271).

§421 bis. Since an expansion of the type (27) may be differenti-

ated term-by-term, it is clear that for solutions of the type (27) the difficulty mentioned in §368, that is the problem of possible spirals, does not arise. But it seems to be quite hard to prove that every solution $\xi_i = \xi_i(t)$ which leads to a simultaneous collision is obtainable by the method of the characteristic exponents s , to be applied in §423 to the existence proof of the particular solutions (27).

§422. If $\xi_i = \xi_i(t)$ is any solution of $m_i \xi_i'' = U_{\xi_i}$ for $0 < t < \text{const.}$, put, as in (18₁), §364 and (15₁), §363,

$$(28_1) \quad t = -\log \tau;$$

$$(28_2) \quad \xi_i = \tau^{-\frac{1}{3}} \xi_i.$$

Then $m_i \xi_i'' = U_{\xi_i}$ is, by (19₁)–(19₂), §364, equivalent to

$$(29_1) \quad \ddot{\xi}_i - \frac{1}{3} \dot{\xi}_i - \frac{2}{9} \xi_i = U_{\xi_i}/m_i; \quad (29_2) \quad U = \sum^* m_j m_k / |\xi_j - \xi_k|,$$

where the dots denote differentiations with respect to t . If $t \rightarrow +0$, then $\tau \rightarrow +\infty$, by (28₁).

Let $\xi_i = \xi_i(t)$ be an homothetic solution of $m_i \xi_i'' = U_{\xi_i}$, and choose this solution so that its energy constant $h = 0$. Such solutions exist, by §378, for arbitrary values of the given m_i , and can be chosen, by §367, either as collinear or as equilateral. In either case they lead, by §378, to a simultaneous collision at some $t = t_0$, say as $t \rightarrow +0$. Since $h = 0$, comparison of §378 with (22₁)–(22₂), §268 shows that $\xi_i(t)$ and t are respectively proportional to the second and third powers of the local uniformizing time parameter u ; so that $\xi_i(t) = t^{\frac{1}{3}} \xi_i(1)$. Denoting the three constant 3-vectors $\xi_i(1)$, which form either an equilateral or a collinear central configuration, by $\alpha_1, \alpha_2, \alpha_3$, one sees from (28₂) that ξ_i is the constant α_i . In other words, $\xi_i(t) \equiv \alpha_i$, together with its consequence $\dot{\xi}_i(t) \equiv 0$, is an equilibrium solution of (29₁) in the sense of §83.

It follows, therefore, from §89 that if $\zeta_i = \zeta_i(t)$, where $i = 1, 2, 3$, denotes a displacement (§86) of this particular solution of (29₁), then the Jacobi equations (§86) which define the ζ_i have constant coefficients. Since $C = 0$, the third components of the six 3-vectors $\xi_i, \dot{\xi}_i$ may be chosen to be $\equiv 0$, by §326. Then the order of the system of the Jacobi equations reduces from $6n = 18$ to $4n = 12$. Thus, the Jacobi equations are of the form (41), §381, where $n = 3$, the prime ($= d/dt$) must be replaced by a dot ($= d/dt$), and $A = (a_{ji})$ is a constant 12-matrix. If s is any of the 12 roots of the equation $\det(sE - A) = 0$ which determines the characteristic exponents (§89), then the Jacobi equations have a solution of the form $\zeta_i = \epsilon \beta_i \exp(st)$, where the vectors β_i are constants and do not all

vanish, while ϵ is an arbitrary constant scalar factor of proportionality, which will be chosen to be positive.

§423. Choosing ϵ small and applying §85–§86, one sees that (29₁) possesses a solution $\xi_i = \xi_i(t)$ which is, on any fixed bounded t -interval, approximated by

$$(30) \quad \alpha_i + \xi_i(t) \equiv \alpha_i + \epsilon \beta_i \exp(st), \quad (i = 1, 2, 3),$$

with an accuracy which increases as ϵ decreases. But $\xi_i(t) \equiv \alpha_i$ is an equilibrium solution of (29₁). Hence, if the fixed characteristic exponent s of the Jacobi equation is negative, an existence theorem on real non-linear analytic differential equations, which is to-day standard,* assures for (29₁) the existence of a family of solutions $\xi_i = \xi_i(t)$ which depends on a small integration constant ϵ and not only is approximated by (30) on a fixed bounded t -interval but has on an infinite interval $\text{Const.} < t < \infty$ a convergent expansion of the form

$$(31) \quad \xi_i = \alpha_i + \beta_i \tau + \sum_{n=2}^{\infty} b_{in}(\epsilon) \tau^n; \quad \tau = \epsilon \exp(st),$$

the real power series in τ having for every fixed small ϵ some region $|\tau| < \text{const.}$ of convergence.

Keep ϵ fixed, define α_{in} for $n = 0, 1, 2, \dots$ by placing $\alpha_{in} = b_{in}(\epsilon) \epsilon^n$ for $n = 2, 3, \dots$, and $\alpha_{i0} = \alpha_i$, $\alpha_{i1} = \beta_i \epsilon$. Then the solution (31) of (29₁) may be written in the form

$$(32) \quad \xi_i = \sum_{n=0}^{\infty} \alpha_{in} \exp(nst); \quad (\text{Const.} < t < +\infty).$$

§424. A solution (32) of (29₁) is, in view of (28₁)–(28₂), equivalent to a solution (27) of the equations $m_i \xi_i'' = U_{\xi_i}$ which (28₁)–(28₂) have transformed into (29₁).

Consequently, the proof of the statements of §421 is complete, provided that at least one of the twelve characteristic exponents s of the Jacobi equations is negative and irrational for suitably chosen values of the masses m_i . In order to verify this proviso, one has to determine the roots s of the equation $\det(sE - A) = 0$. This is a tiresome task of the same elementary type as the corresponding calculation described in §381.

* Poincaré's doctoral thesis centers about this theorem. The existence of the expansions in question was suggested to him by a remark of Darboux.

On carrying out these elementary calculations for the present case, one finds that, if the equilibrium solution $\xi_i \equiv \alpha_i$ of (29₁) is chosen as equilateral, eight of the roots of the equation $\det(sE - A) = 0$ of degree twelve are, as in §382, of a trivial type. If one removes these trivial roots, the resulting biquadratic equation is easily solvable. One of its roots $s = s(m_1, m_2, m_3)$ is found to be such as to be negative for arbitrary (m_1, m_2, m_3) and to depend on the masses m_i (which occur in the coefficients). Finally, this $s = s(m_1, m_2, m_3)$ does attain irrational values, since it is an algebraic, and hence continuous, function of the m_i .

The situation is similar if the underlying central configuration is collinear, instead of being equilateral.

§425. It is clear that the method of §421–§424 may be extended to the case of the simultaneous collision (§335) of $n \neq 3$ bodies, and that the same holds for the method of §415–§420 in case of a binary collision (§349); the identical substitution (14₂) belonging to $n = 3$ being, for $n > 3$, replaced by $n - 2$ identical substitutions.

More generally, suppose that there exist in the barycentric inertial coordinate system ξ less than n points, say $O_1, \dots, O_l, \dots, O_q$, such that, as $t \rightarrow +0$, exactly n_l of the n bodies m_i tend to the position O_l , where every $n_l \geq 1$ and $n_1 + \dots + n_q = n > q (\geq 1)$; so that at least one $n_l \geq 2$. Then it is easy to see that for those l for which $n_l > 2$ the considerations of §361–§368 and §421–§424 may be extended to the group of the n_l bodies which collide at O_l ; and that for those l for which $n_l = 2$ the considerations of §349–§350 and §415–§420 may be extended to the binary group which collides at O_l .

The trouble is (§411) that, unless $n = 3$ (§412–§413), it is not known whether or not there must exist points O_1, \dots, O_q whenever the singularity condition $r(t) \rightarrow 0$ of §409–§410 is satisfied.

The Problem of Three Bodies

§426. Let a collision in the problem of $n = 3$ bodies be called *continuable* if it is not a simultaneous collision of the type (ii), §421, i.e., if it does not lead to a transcendental singularity; so that, in particular, every binary collision is continuable.

Consider any fixed solution $\xi_i = \xi_i(t)$ of the problem of $n = 3$ bodies. Starting at any initial $t = t_0$ at which all three $\rho_{jk} > 0$, follow the motion for $t < t_0$, say. Then there are three cases possible: either

(I) one does not arrive at a t at which (1), §407 vanishes, in which case §409 shows that the motion proceeds without singularities till $t = -\infty$;

or else one is led, by §413, to a collision, in which case either

(II) one arrives at the date of a continuable collision or

(III) one arrives at the date of a non-continuable collision.

Suppose that (II) is the case, and let $t^{(1)}$ denote the date of collision which follows the initial date t_0 . Then, on considering the analytic continuation of the given motion $\xi_i = \xi_i(t)$ beyond $t^{(1)}$, there may arise for $t < t^{(1)}$ any of the three cases (I), (II), (III). Hence, if one proceeds as before and repeats this process as long as it can be repeated, there clearly results an alternative, to the effect that

either one is led, after a finite number (≥ 0) of cases (II), to one of the two cases (I), (III), so that the successive analytic continuations of the given motion $\xi_i = \xi_i(t)$ for $t < t_0$ are obtained in a finite number of steps;

or else one never arrives at either of the two cases (I), (III), so that the process of the analytic continuation through successive dates of collisions has to be repeated infinitely often, thus leading to an infinite sequence $t^{(1)} > t^{(2)} > \dots > t^{(m)} > \dots$ of continuable collisions.

But it will now be shown that in the latter case $t^{(m)} \rightarrow -\infty$ as $m \rightarrow \infty$.

§427. Suppose, if possible, that the infinite sequence $t^{(1)}, t^{(2)}, \dots$ exists and does not tend to $-\infty$. Then, since $t^{(m)} > t^{(m+1)}$, there exists a finite t^* such that $t^{(m)} \rightarrow t^*$ as $m \rightarrow \infty$. Choose the origin of the t -axis so that $t^* = 0$. Let the signs \lim ; $\overline{\lim}$ refer to the limit process $\lim t = +0$, where t varies continuously, passing, in particular, through all the discrete collision dates $t^{(m)}$ which cluster at $+0$.

In this sense, one has $\lim r(t) = 0$, where $r = \min(\rho_{12}, \rho_{23}, \rho_{31})$. In fact, the assumption $\overline{\lim} r(t) > 0$ implies the existence of a sequence t_1, t_2, \dots such that t_m lies between $t^{(m)}$ and $t^{(m+1)}$, while $r(t_m)$ exceeds for every m a fixed positive number, say r^* . Then, since $t_m \rightarrow +0$, one can apply (4₂) in the same way as in §409, thereby obtaining the same contradiction as in §409.

Since $\lim r(t) = 0$, the reasoning of §411 bis holds without change. Thus, $\lim J''(t) = +\infty$; and there exists a non-negative $\lim J(t) \leq +\infty$.

Next, the existence of $\lim J(t) \leq +\infty$ implies that $\lim r(t) = 0$ may be replaced by the sharper statement that $\lim \rho_{jk}(t) = 0$ holds

for at least one of the three ρ_{jk} . For otherwise one could select a sequence of dates t_1, t_2, \dots such that t_m satisfies the same condition as in §412 and lies between $t^{(m)}$ and $t^{(m+1)}$. And this leads to the same contradiction as in §412.

Consequently, one can choose the notations so that $\lim \rho_{12}(t) = 0$ as $\lim t = +0$.

§428. By the definition of the $t^{(m)}$, at least one of the three $\rho_{jk}(t)$ vanishes at every fixed $t^{(m)}$, where $t^{(m)} \rightarrow +0$ as $m \rightarrow \infty$. Whether the collision is binary or simultaneous at a fixed $t^{(m)}$, the last remark of §414 bis shows that, although $J''(t)$ becomes (positively) infinite, $J'(t)$ remains continuous at every $t^{(m)}$. Furthermore, $\lim J''(t) = +\infty$ implies that $J'(t)$, hence also $J(t)$ itself, is monotone in a sufficiently small neighborhood $0 < t < \epsilon$ of $\lim t = +0$. Since $J(t) > 0$ between $t = t^{(m)}$ and $t = t^{(m+1)}$, and since $t^{(m)}$ tends to $+0$ as $m \rightarrow \infty$, it follows that $J(t^{(m)}) \neq 0$ for every sufficiently large m . In other words, the collision which takes place at $t^{(m)}$ is a binary collision from a certain m onward.

Since $\lim \rho_{12}(t) = 0$, it follows that either all three $\lim \rho_{jk}(t) = 0$ or one and the same ρ_{jk} , namely ρ_{12} , vanishes at $t^{(m)}$, when m varies and is sufficiently large. But it will now be shown that either of these cases, which are not mutually exclusive, leads to a contradiction. These contradictions will disprove the existence of the finite t^* , assumed at the beginning of §427.

§429. Suppose first that all three $\lim \rho_{jk}(t) = 0$ as $\lim t = +0$. Then, although the dates $t^{(m)}$ of the collisions cluster at $\lim t = +0$, nothing hinders a repetition of the considerations of §335–§338 bis, the necessary modifications being of an obvious nature in view of the fact that the intermediary collisions are all binary collisions between m_1 and m_2 (§428). Thus, (18₁), §337 is applicable to $\lim t = +0$; so that $\lim J''(t)\sqrt{J(t)}$ exists and is distinct from 0. But $J(t)$ cannot vanish for t sufficiently close to $\lim t = +0$, since the collision at $t^{(m)}$ is not a simultaneous collision from a certain m onward. Consequently, $J''(t)$ is finite for every t sufficiently close to $\lim t = +0$. This is a contradiction, since, as pointed out in §428, one has $J''(t^{(m)}) = +\infty$ for every m ; while $t^{(m)} \rightarrow +0$ as $m \rightarrow \infty$. This disposes of the first of the two cases found at the end of §428.

In order to disprove the possibility of the second case, it may clearly be assumed that, while $\lim \rho_{12}(t) = 0$ but not all three $\lim \rho_{jk}(t) = 0$, one has $\rho_{12}(t^{(m)}) = 0$ for every sufficiently large m .

On the other hand, the same proof as in §412 shows that there exists a common non-vanishing $\lim \rho_{13}(t) = \lim \rho_{23}(t) \leq +\infty$. Hence, if t is sufficiently close to $\lim t = +0$, both $\rho_{j3}(t)$ exceed a fixed positive lower bound; and so nothing hinders a repetition of the considerations of §349 bis, the necessary modification being of an obvious nature in view of the fact that the intermediary collisions are all binary collisions between m_1 and m_2 (§428). Thus, (8_2) , §413 is applicable for $\lim t = +0$. But it is clear from the remarks of §336–§337 that (8_2) , §413 implies (8_1) , §413. And (8_1) , §413 shows that $t^3 \rho_{12}(t)$, hence also $\rho_{12}(t)$, is positive for every sufficiently small $t > 0$. Consequently, $\rho_{12}(t^{(m)}) = 0$ cannot hold for every sufficiently large m .

This contradiction completes the proof of the fact announced at the end of §426.

§430. Thus, the dates of continuable collisions (§426) cannot cluster at a finite limiting t^* . It follows, therefore, from the alternative formulated before the last statement of §426, that if a solution $\xi_i = \xi_i(t)$ of the problem of $n = 3$ bodies cannot be continued analytically till $t = -\infty$ (or $t = +\infty$), the solution can cease to exist only at a finite t which is an isolated transcendental (logarithmic) singularity and represents the second of the two cases (i), (ii) of §421.

On comparing this with §413, one sees, in particular, that if the solution has an invariable plane (e.g., if the solution is not planar), then the solution exists from $t = -\infty$ to $t = +\infty$, provided that the motion is thought of as continued analytically through all the dates of binary collisions; it being understood that the number of such dates may be finite (≥ 0) or infinite.

Actually, the example mentioned at the end of §346 bis shows that a solution which has no invariable plane may also be such as to lead to no simultaneous collision at all. Finally, on choosing $h < 0$ in the homographic solution used, at the beginning of §421, to prove the statement (i), one sees that a solution may exist from $t = -\infty$ to $t = +\infty$ even when it leads to infinitely many simultaneous collisions.

§431. It should be mentioned that if there exists an invariable plane, i.e., if $C \neq 0$, then not only is $J = \mu^{-1} \sum^* m_j m_k \rho_{jk}^2$ positive (or, what is the same thing, $\text{Min}(\rho_{12}, \rho_{23}, \rho_{32}) \equiv r > 0$) for every t (§335) but also one has $\lim J > 0$, i.e., $\lim r > 0$, as $t \rightarrow \pm \infty$. The proof

of this theorem, which is based on the inequalities of §333–§334 bis, is at present too lengthy to be reproduced here.*

§431 bis. Usually, the proof of the particular case $C \neq 0$ of the fact which was formulated at the end of §426 (and proved, for both cases $C \neq 0$, $C = 0$, in §427–§429) is based on the theorem of §431. Notice, however, that the theorem of §431 is not applicable to those solutions with $C = 0$ which possess only binary collisions (or, perhaps, no collisions at all) for $-\infty < t < +\infty$.

§432. It is clear from the fact formulated at the end of §426 that if a solution of the problem of three bodies does not possess a non-continuable singularity (e.g., if $C \neq 0$), then the regularizing time variable u , when defined by (26), §420 bis, tends monotonously to $\pm\infty$ as $t \rightarrow \pm\infty$.

In the particular case $C \neq 0$, the theorem of §431 supplies additional information. In fact, it then readily follows from the footnote to §408 by direct analytic continuation along the real u -axis, that the three barycentric position vectors ξ_i , when considered as functions of the time variable u , are regular analytic in a strip $|\operatorname{Re}(u\sqrt{-1})| < \text{const.}$ about the real axis of the complex u -plane, $\operatorname{Re}(z)$ denoting the real part of z .

§432 bis. If this strip $|\operatorname{Re}(u\sqrt{-1})| < \text{const.}$ is mapped in a one-to-one and conformal manner on the interior of the unit circle of a complex w -plane,† the ξ_i may, of course, be developed into regular power series in w which are convergent for $|w| < 1$; so that, in virtue of the transformations $w = w(u)$ and $u = u(t)$, there result for the $\xi_i = \xi_i(t)$ certain expansions which are valid for $-\infty < t < +\infty$. This trivial restatement of the purely function-theoretical result of §432 is often given undue emphasis by saying that, if $C \neq 0$, the problem of three bodies is solved, since the ξ_i can be developed into series.

Incidentally, it turns out that the expansions in question are convergent so slowly as to be, for all practical purposes, completely useless even in so simple a case as an equilateral homothetic solution.

§433. It is clear from §430 that the solutions of the problem of three bodies are, in general (e.g., whenever $C \neq 0$), unrestricted solu-

* Only the case $h < 0$ is awkward, since if $h \geq 0$, the theorem readily follows by the simple method of §332–§332 bis.

† The explicit form of such a mapping is $w = (e^u - 1)/(e^u + 1)$, if $\text{const.} = \frac{1}{2}\pi$.

tions in the sense of §119. Thus there arises the question, what do all the results obtained actually mean from the point of view of the "problem of integration" of the equations of motion. In order to formulate an answer to this question, it will be necessary to return to the elimination of the linear momentum (centre of mass) and of the angular momentum.

§434. The reduction of (9₁), §384 to (32), §394 has used the conservation of the linear and angular momenta, but not the conservation of the energy. Correspondingly, (33), §394 contains the angular momentum constant $|C|$ but not the energy constant h . By using the energy integral $H = h$ also, where H is given by (33), §394, one of the 8 variables $I, P_i; \iota, \rho_i$ may be eliminated; so that the system (32), §394 of order 8 reduces to a system of the form $z'_k = Z_k(z_1, \dots, z_7); k = 1, \dots, 7$, where the known functions Z_k of the z_k depend on both constants $|C|, h$. Since this system of order 7 does not contain t explicitly, it may be replaced by a system of order 6 which contains the independent variable; the latter being one of the z_k on the assumption that not all $z_k(t) = \text{const.}$ Actually, this non-conservative system of order 6 appears in the form of a non-conservative Hamiltonian system with $6:2 = 3$ degrees of freedom, if one applies to (32), §394 the method of §181.

§435. For instance, if $\iota = \iota(t)$ is not independent of t along the solution under consideration, then, by (18), §181,

$$\begin{aligned} \dot{P}_i &= -K_{\rho_i}, & \dot{\rho}_i &= K_{P_i}; & i &= 1, 2, 3, \\ K &= K(P_1, P_2, P_3, \rho_1, \rho_2, \rho_3; h, |C|), \end{aligned}$$

where the dots denote differentiations with respect to the time variable ι . This (non-conservative) Hamiltonian system with 3 degrees of freedom is an intrinsic representation of the problem of $n = 3$ bodies, since the coordinates are the mutual distances ρ_i , while the time variable is the inclination of the varying plane $\Pi(t)$ of the three bodies towards the fixed plane Π^* , a plane defined in §394 in an intrinsic manner.

§436. It remains to determine those solutions of the problem of $n = 3$ bodies for which the assumption $\iota(t) \neq \text{const.}$ of the intrinsic Hamiltonian equations of §435 is violated. It is certainly violated if the solution is planar, since then $\iota(t) = \text{const.} = 0$. However, planar solutions may be disregarded, since for these §399 supplies

a Hamiltonian system of the same form, with t instead of ι as independent variable (and with a Hamiltonian function which is conservative). Unfortunately, $\iota(t) = \text{const.}$ is possible for certain non-planar solutions also. In fact, it is easily verified from §346 that the inclination ι has the constant value $\frac{1}{2}\pi$ for either type (i)–(ii) of non-planar isosceles solutions. As far as present knowledge goes, it is possible that no further exceptions to §435 exist. Actually, the enumeration of all solutions with $\iota(t) = \text{const.}$ seems to be an intricate question (although the answer may be trivial); it might depend on function-theoretical considerations of the type indicated in §389. At any rate, it is not obvious at all that const. cannot be distinct from 0 and $\frac{1}{2}\pi$, and that $\text{const.} = \frac{1}{2}\pi$ is possible for isosceles solutions only.

§437. For a fixed value of $|C|$ in (33) and for a fixed energy h , let $\mathbf{M}_7 = \mathbf{M}_7(|C|; h)$ denote the manifold (or, more correctly, point-set of generic local dimension number 7) which results from the 8-dimensional phase space of (32), §394 on isoenergetic reduction.

More precisely, let \mathbf{M}_7 be the locus of those points in the *admissible* $(I, \iota, \dots, P_3, \rho_3)$ -region on which the function (33), §394 attains the fixed value h , where the italicized proviso has the rôle of subjecting the topology of \mathbf{M}_7 to appropriate requirements. For instance, the inclination ι must be thought of as an angular variable (mod π), while the subspace of the 3 distances ρ_i ought to be defined by the inequalities $0 < \rho_i < \rho_j + \rho_k$, if Δ were (as it was at the beginning of §394) required to be a non-degenerate triangle. Actually, the complete manifold of all possible states of motion of the problem of three bodies is obtained only if one also includes, on the one hand, the limiting cases of syzygies and collinear solutions, where $|\Delta| = 0 < \rho_i = \rho_j + \rho_k$ for one (i, j, k) , and, on the other hand, the limiting cases of binary and general collisions, where at least one $\rho_i = 0$. In fact, §498–§500 will show in a relatively simple case, how fundamental are the collisions for the understanding of the topological structure. Of course, it can be decided only by detailed discussions, what is *admissible* for $(I, \iota, P_1, P_2, P_3)$ when (ρ_1, ρ_2, ρ_3) is in any of the limiting cases.

§438. All these remarks are to the effect that the topology of $\mathbf{M}_7 = \mathbf{M}_7(|C|; h)$ is thought of as being identical with the topology of all those states of the reduced problem of three bodies which are compatible with the given values of the constants $|C|; h$, constants

conserved along every solution path of (9₁), §384. This implies that, from the topological point of view, \mathbf{M}_7 is, for fixed $|C|; h$, intrinsically connected with the problem of three bodies (so that, in particular, \mathbf{M}_7 is independent of the choice of the phase variables and may, therefore, be defined by means of (10₁)–(10₃), §384 and $H(\eta_1, \dots, \xi_3) = h$ also). Thus, a description of \mathbf{M}_7 might become of fundamental importance (cf. §227). Unfortunately, nothing explicit is known as to the topological structure of \mathbf{M}_7 .

§439. It is easy to show that, barring the lower-dimensional limiting case of collinear solutions, the manifold $\mathbf{M}_7(|C|; h)$ does not contain any solution path consisting of singular and only singular points of this manifold, provided that $|C|; h$ do not satisfy the condition $1 + h^0|C^0|^2 = 0$ [mentioned at the end of §378, where $|C^0|; h^0$ are defined in terms of $|C|; h$ (and m_1, m_2, m_3) by means of the formulae of §375 and §378]. On the other hand, in the case of those $|C|; h$ which satisfy the condition $1 + h^0|C^0|^2 = 0$ and determine, therefore, equilateral triangles of relative equilibrium, there correspond to these equilibrium solutions single points (instead of curves) of the respective manifolds $\mathbf{M}_7(|C|; h)$; and these isolated points turn out to be singular points of the latter.

The proof proceeds as follows: Since the singularities $\rho_i = 0$ and $\Delta = 0$ of the function (33), §394 need not be considered, it is clear from the footnote to §394 that the function (33), §394 may be assumed to be regular analytic along the exceptional solutions under consideration. But then the manifold \mathbf{M}_7 , which has been defined by the equation $H = \text{const.} = h$, cannot become singular at a point at which the partial derivatives of the first order of the function (33), §394 of eight variables do not vanish simultaneously. Hence, (32), §394 shows that all eight phase variables I, \dots, ρ_3 must be independent of t along the exceptional solutions in question. Since, in particular, the ρ_i are independent of t and belong, therefore, to a solution of relative equilibrium, it follows from §367 and from the exclusion of collinear solutions, that the three constants $\rho_i = \rho$ determine an equilateral triangle. This fact, when combined with (32)–(33), §394 and (ii), §371, readily shows that also $\iota; I, P_i$ are independent of t and determine, together with the $\rho_i = \rho$, a singular point of \mathbf{M}_7 ; so that the proof is complete.

§440. As seen from §200–§201, every new generation usually is compelled to reinterpret what *the* “problem” of three bodies actually

is. Until Birkhoff realized and further developed Poincaré's geometrical ideas concerning dynamical systems with two degrees of freedom, the answer to the question used to be this: On the one hand, the problem of three bodies cannot be "solved," in view of the established non-existence of integrals of specific type (§129, §320 bis); while, on the other hand, the problem of three bodies may be considered as "solved," in view of the convergence of certain expansions, established along the whole t -axis (§432 bis). To-day one is inclined to consider the first of these statements as inadequate, and the second as quite meaningless, and accordingly to formulate the problem of three bodies in terms of an "incompressible flow" on a seven-dimensional manifold, as follows:

For a fixed pair of values of the conservation constants $|C|, h$, consider all those solutions $\xi_i = \xi_i(t)$ of the problem of three bodies which are continuable for $-\infty < t < +\infty$, where it is understood that the latter restriction is necessary only when $C = 0$. Whether $C = 0$ or $C \neq 0$, the reduced state of the solution $\xi_i = \xi_i(t)$; $i = 1, 2, 3$ at a fixed t is represented by a point of the manifold $\mathbf{M}_7 = \mathbf{M}_7(|C|; h)$. Thus, the whole solution $\xi_i = \xi_i(t)$; $-\infty < t < +\infty$, ($i = 1, 2, 3$), is represented on $\mathbf{M}_7 = \mathbf{M}_7(|C|; h)$ by a path which degenerates into a point only in case $\xi_i = \xi_i(t)$ is a solution of relative equilibrium. These ∞^7 paths, which do not intersect each other, determine on $\mathbf{M}_7 = \mathbf{M}_7(|C|; h)$ a transformation group τ^t , $-\infty < t < +\infty$, of the type described in §121 (at least if the case $C = 0$ compatible with an unrestricted solution, or rather any such solution on $\mathbf{M}_7(0; h)$, is excluded). It is easy to verify that the "flow" of the paths which is defined on $\mathbf{M}_7(|C|; h)$ by the transformation group $\tau^t = \tau^t(|C|; h)$, $-\infty < t < +\infty$, is incompressible in the sense of §122, if the topological manifold \mathbf{M}_7 is thought of as embedded into a canonical phase space (e.g., into the $(1, \dots, \rho_3)$ -space of (32)–(33), §394). And the problem of three bodies requires, for arbitrarily fixed $(|C|; h)$, the topological investigation of this flow.

CHAPTER VI

INTRODUCTION TO THE RESTRICTED PROBLEM

The restricted problem of three bodies.	§441–§445
Regularization.	§446–§461
The syzygical potential curve.	§462–§468
The potential surface.	§469–§477
The non-planar restricted problem.	§478–§488
Lunar systems.	§489–§502
Periodic lunar orbits.	§503–§515
Lunar theory.	§516–§529

The Restricted Problem of Three Bodies

§441. Let P_1, P_2 denote the two particles in the problem of $n = 2$ bodies. Let the total mass be the unit of mass; so that the mass of P_1 is $1 - \mu$, if μ denotes the mass of P_2 . Thus, §343 (in conjunction with §207) shows that the equations of motion are given by (2₁), §241, where x, y denote the Cartesian coordinates of P_2 in an (x, y) -plane which contains P_2 for every t , has P_1 as origin, and possesses coordinate axes which are parallel to those of an inertial coordinate system.

Suppose, in particular, that the integration constants determine the motion of P_2 about P_1 as a circular path. Choose the unit of length to be the radius of this circle. Then §276 shows that P_2 has in the (x, y) -plane a constant angular velocity, n , and that $n^2 \cdot 1^3 = 1$; so that, by the end of §214, one can choose $n = +1$ without loss of generality. Thus, the coordinates (x, y) of P_2 at an arbitrary date t are $(\cos t, \sin t)$, if the direction of the positively oriented x -axis is chosen so as to point towards that position of P_2 which belongs to $t = 0$.

Now consider a third particle, P , which moves in the (x, y) -plane in such a way that, while it is subject to the Newtonian attractions of P_1 and P_2 , it does not disturb the Keplerian motion of the two bodies P_1, P_2 . Although this assumption is at variance with Newton's law of gravitation, it gives a reasonable approximation to the actual situation in case the mass of the "infinitesimal" body P is much smaller than the mass of either of the "finite" bodies P_1, P_2 . The resulting model is called the restricted problem of three bodies.

It is a problem with two degrees of freedom. In fact, one sees

from (11₁)-(11₂), §342 that if \bar{x} , \bar{y} denote the coordinates x , y of P , then

$$\bar{x}'' + (1 - \mu + 0) \cdot \frac{\bar{x}}{(\bar{x}^2 + \bar{y}^2)^{\frac{3}{2}}} = \Omega_{\bar{x}},$$

$$\bar{y}'' + (1 - \mu + 0) \cdot \frac{\bar{y}}{(\bar{x}^2 + \bar{y}^2)^{\frac{3}{2}}} = \Omega_{\bar{y}},$$

$$\Omega = \mu \cdot \left(\frac{1}{|(\bar{x} - \cos t)^2 + (\bar{y} - \sin t)^2|^{\frac{3}{2}}} - \frac{\bar{x} \cos t + \bar{y} \sin t}{|(\cos t)^2 + (\sin t)^2|^{\frac{3}{2}}} \right),$$

since 0 , $1 - \mu$, μ are the masses, and (\bar{x}, \bar{y}) , $(0, 0)$, $(\cos t, \sin t)$ the coordinates, of P , P_1 , P_2 , respectively. Clearly, one can write these equations in the form $\bar{x}'' = \bar{U}_{\bar{x}}$, $\bar{y}'' = \bar{U}_{\bar{y}}$, where \bar{U} denotes the (non-conservative) force function $\bar{U} \equiv \bar{U}(\bar{x}, \bar{y}; t) = (1 - \mu)/(\bar{x}^2 + \bar{y}^2)^{\frac{1}{2}} + \Omega$. Accordingly, the equations of motion of P have the non-conservative Lagrangian function \bar{L} defined by

$$(1_1) \quad \bar{L} = \frac{1}{2}(\bar{x}'^2 + \bar{y}'^2) + \bar{U}; \quad (1_2) \quad \bar{U} = (\bar{x}^2 + \bar{y}^2)^{-\frac{1}{2}} + \mu \bar{F}(\bar{x}, \bar{y}; t);$$

$$(1_3) \quad \bar{F} = ((\bar{x} - \cos t)^2 + (\bar{y} - \sin t)^2)^{-\frac{1}{2}} \\ - (\bar{x}^2 + \bar{y}^2)^{-\frac{1}{2}} - (\bar{x} \cos t + \bar{y} \sin t).$$

§442. The restricted problem of three bodies was first considered by Euler in connection with one of his lunar theories. The mathematical and astronomical significance of this model was, however, understood only much later.

First, Jacobi observed that the problem is, as a matter of fact, a conservative problem with two degrees of freedom. In order to see this, it is sufficient to replace (\bar{x}, \bar{y}) by a coordinate system (ξ, η) which rotates about the common origin, P_1 , so as to transform P_2 to rest; so that

$$(2) \quad \xi = \bar{x} \cos t + \bar{y} \sin t, \quad \eta = -\bar{x} \sin t + \bar{y} \cos t,$$

the coordinates $(\cos t, \sin t)$ of P_2 thus being transformed into $(1, 0)$ for every t . Substitution of the inverse of (2) into (1₁)-(1₃) readily shows that if one puts $\bar{L} \equiv L$ in accordance with §95, then

$$(3_1) \quad L = \frac{1}{2}(\xi'^2 + \eta'^2) + (\xi\eta' - \eta\xi') + \left\{ \frac{1}{2}(\xi^2 + \eta^2) + U \right\};$$

$$(3_2) \quad U = (\xi^2 + \eta^2)^{-\frac{1}{2}} + \mu F;$$

$$(3_3) \quad F = ((\xi - 1)^2 + \eta^2)^{-\frac{1}{2}} + (\xi^2 + \eta^2)^{-\frac{1}{2}} - \xi.$$

And (3₂)–(3₃) imply that the Lagrangian function (3₁) is, while irreversible, conservative. It follows, therefore, from §155 that the Lagrangian equations $[L]_{\xi} = 0$, $[L]_{\eta} = 0$ of P admit the integral $\frac{1}{2}(\xi'^2 + \eta'^2) - \{ \quad \} = \text{const.}$ This integral, which is called the integral of Jacobi, expresses the conservation of relative energy, the term $\frac{1}{2}(\xi'^2 + \eta'^2)$ of $\{ \quad \}$ representing the force function of the centrifugal forces, which are introduced by the uniform rotation (2); while the term $(\xi\eta' - \eta\xi')$ of (3₁) corresponds to Coriolis forces, which do not appear in the energy (cf. §155).

This conservative formulation of the restricted problem of three bodies became fundamental, first in Delaunay's elaborate lunar theory, and then, under its apparent influence, during the last quarter of the 19th century. On the one hand, G. W. Hill developed at that time his lunar theory, which is based on (3₁)–(3₂) and, as elaborated in its details by E. W. Brown, is to-day the most precise treatment of a problem ever dealt with in celestial mechanics (precision being meant in both the theoretical and the numerical sense of the word). On the other hand, it turned out that the model of the restricted problem of three bodies yields a tolerable approximation in many cases of minor planets also.

At the same time, this model aroused the interest of Poincaré, whose mathematical work in dynamics centered about it. In fact, he considered (3₁)–(3₂) as a prototype of those dynamical problems which have two degrees of freedom and are not "integrable" in the sense in which a problem with a single degree of freedom is. In one respect, the irreversible problem (3₁)–(3₂) is more complicated than the simplest "non-integrable" system, the latter being reversible (and having two degrees of freedom). Actually, there are some indications to the effect that the topology of the restricted problem of three bodies, and therefore also this problem itself, is, though difficult enough, too simple to be characteristic of a "generic" dynamical system with two degrees of freedom. At any rate, almost everything mathematically significant in the progress of general analytical mechanics during the 20th century, and in particular the dynamical work both of Levi-Civita and of Birkhoff, was originally directed towards, when not influenced by, an investigation of the restricted problem of three bodies.

Incidentally, the restricted problem of three bodies often (though not always) indicated what to expect in the problem of $n = 3$ bodies proper. For instance, the regularization (Sundman; Levi-Civita) of

the latter problem in case of binary collisions (§415–§420) was preceded by the regularization (Thiele and Burrau; Levi-Civita) of the restricted problem (§446–§452).

§443. According to §442, the bodies P_1, P_2 rest at the respective points $(0, 0), (1, 0)$ of the rotating coordinate system (ξ, η) ; so that the centre of mass rests at $(\mu, 0)$, the masses of P_1, P_2 being $1 - \mu, \mu$, respectively. It will be convenient to replace (ξ, η) by a new* coordinate system, (x, y) , which is barycentric; so that

$$(4) \quad \xi = x + \mu, \quad \eta = y,$$

$(-\mu, 0)$ and $(1 - \mu, 0)$ being the new coordinates of P_1 and P_2 for every t . Thus, (x, y) is a coordinate system which rotates uniformly about the centre of mass of P_1 and P_2 .

Substitution of (4) into (3₁)–(3₃) readily gives

$$(5_1) \quad L = \frac{1}{2}(x'^2 + y'^2) + (xy' - yx') + U(x, y);$$

$$(5_2) \quad U = \frac{1}{2}(x^2 + y^2) + \frac{1 - \mu}{|(x + \mu)^2 + y^2|^{\frac{1}{2}}} + \frac{\mu}{|(x - 1 + \mu)^2 + y^2|^{\frac{1}{2}}},$$

if one omits the additive terms $\mu\eta'$ and $\frac{1}{2}\mu^2$. This omission is justified by §156, since $\mu\eta'$ is the derivative G' of $G \equiv \mu\eta$, while $\frac{1}{2}\mu^2 = \text{const.}$

For reasons which will become apparent later (cf. §517), the rotating barycentric coordinate system (x, y) is called the synodical coordinate system. If $\mu = 0$, then (5₁)–(5₂) reduce to (5₁), §300; so that the present terminology is the same as in the limiting case of §300.

The x -axis of the synodical coordinate system is called the axis of syzygies. This terminology agrees with that of §327, since the first two of the three bodies $P_1, P_2; P$ rest on the x -axis.

According to (5₁)–(5₂), the Lagrangian equations $[L]_x = 0, [L]_y = 0$ and their energy integral may be written as

$$(6_1) \quad x'' - 2y' = U_x, \quad y'' + 2x' = U_y;$$

$$(6_2) \quad x'^2 + y'^2 = 2U(x, y) - C,$$

if $-\frac{1}{2}C$ denotes (as in the limiting case $\mu = 0$ of §300) the energy constant; C itself is called the Jacobi constant.

* This coordinate system (x, y) must not be confused with the coordinate system (x, y) of §441 which now is denoted by (\bar{x}, \bar{y}) .

If X, Y denote the momenta and $H(X, Y; x, y)$ is the Hamiltonian function belonging to (5₁)–(5₂), then, according to §229,

$$(7_1) \quad X = x' - y, \quad Y = y' + x;$$

$$(7_2) \quad H = \frac{1}{2}(X^2 + Y^2) - (xY - yX) - V(x, y);$$

$$(7_3) \quad V(x, y) = U(x, y) - \frac{1}{2}(x^2 + y^2);$$

$$(7_4) \quad H = h; \quad (7_5) \quad h = -\frac{1}{2}C.$$

§443 bis. In terms of the bipolar coordinates (33), §56, one can write the force function (5₂), §443 for two arbitrary masses $\mu, 1 - \mu$ in the symmetric form

$$U = (1 - \mu) \cdot (\frac{1}{2}r_1^2 + r_1^{-1}) + \mu \cdot (\frac{1}{2}r_2^2 + r_2^{-1}) + \text{const.};$$

$$\text{const.} = -\frac{1}{2}\mu(1 - \mu), \text{ since } (1 - \mu)r_1^2 + \mu r_2^2 \equiv x^2 + y^2 + \mu(1 - \mu).$$

§444. If the last sum, which is introduced by the centrifugal forces, were missing, U would reduce to $U = (1 - \mu)/r_1 + \mu/r_2$; so that, if also the Coriolis forces, represented in (5₁) by $(xy' - yx')$, were missing, the problem would reduce to the elementary problem of §203, which can be solved in terms of elliptic functions.

§444 bis. It may be mentioned that if the two masses are equal, then it is necessary to disregard only the Coriolis, and not also the centrifugal, forces, in order to obtain a problem solvable by quadratures (leading again to elliptic functions). For if $1 - \mu = \mu$, then $U = \frac{1}{4}(r_1^2 + r_2^2) + \frac{1}{2}(r_1^{-1} + r_2^{-1}) + \text{const.}$, by §443 bis; so that the reversible problem belonging to the Lagrangian function $L = \frac{1}{2}(x'^2 + y'^2) + U$ is easily seen to become of the type considered in §194, if the variables are chosen in the same way as in §203.

§445. The energy integral (6₂) is the only “known” integral of (6₁). In fact, the negative results mentioned in §320 bis for the problem of $n(\geq 3)$ bodies can be established for the restricted problem (6₁) also, the single integral (6₂) playing the rôle of the group of all ten conservation integrals (§320). However, these negative results concerning (6₁) are not of a definitive nature, since the remarks of §320 bis hold again.

Regularization

§446. The Lagrangian function (5₁), §443 is of the form (5₁), §229, with $f(x, y) \equiv 1$; so that $\omega \equiv 1$, by (3), §228. Thus, on applying

(11₂)–(13₂), §230 to an arbitrary conformal mapping $x + iy \equiv z = z(\zeta) \equiv z(\xi + i\eta)$, one obtains*

$$(8_1) \quad \ddot{\xi} - 2|z_\zeta|^2 \dot{\eta} = \overline{U}_\xi, \quad \ddot{\eta} + 2|z_\zeta|^2 \dot{\xi} = \overline{U}_\eta;$$

$$(8_2) \quad \bar{t}' = 1/|z_\zeta|^2, \quad (z_\zeta \neq 0),$$

where the dots refer to the time variable $\bar{t} = \bar{t}(t)$ which follows from (8₂) by a quadrature, and

$$(9_1) \quad \frac{1}{2} (\dot{\xi}^2 + \dot{\eta}^2) - \overline{U} = 0;$$

$$(9_2) \quad \overline{U} \equiv \overline{U}(\xi, \eta; -\frac{1}{2}C) = |z_\zeta|^2(U - \frac{1}{2}C).$$

§447. If neither $\mu = 0$ nor $\mu = 1$, the real finite singularities of the analytic force function (5₂), hence also those of the differential equations (6₁), are seen to be the points $(x, y) = (1 - \mu, 0)$ and $(x, y) = (-\mu, 0)$, at which the two attracting masses $\mu, 1 - \mu$ rest. If $\mu = 0$, the first of these singularities disappears, while the second is the one which, in §268–§269, was regularized by the transformation $z = \zeta^2$ of §259. This suggests that if $0 < \mu < 1$, the second and the first of the singularities may be regularized by choosing $z = -\mu + \zeta^2$ and $z = (1 - \mu) + \zeta^2$, respectively.

For reasons of symmetry, it will be sufficient to consider the singularity at $(x, y) = (-\mu, 0)$; so that the mapping is $z = -\mu + \zeta^2$, i.e., the mapping $x = -\mu + \xi^2 - \eta^2, y = 2\xi\eta$ considered in §54. Thus, (8₁)–(8₂) may be written as

$$(10_1) \quad \ddot{\xi} - 8(\xi^2 + \eta^2)\dot{\eta} = \overline{U}_\xi, \quad \ddot{\eta} + 8(\xi^2 + \eta^2)\dot{\xi} = \overline{U}_\eta;$$

$$(10_2) \quad \bar{t} = 4(\xi^2 + \eta^2);$$

while (5₂) shows that (9₂) becomes

$$(11) \quad \overline{U} = 4(\xi^2 + \eta^2) \left(\mu^2 - 2(\xi^2 - \eta^2)\mu + (\xi^2 + \eta^2)^2 + \frac{1 - \mu}{\xi^2 + \eta^2} + \frac{\mu}{\{1 - 2(\xi^2 - \eta^2) + (\xi^2 + \eta^2)^2\}^{\frac{1}{2}}} - \frac{1}{2}C \right).$$

It is clear from (11) that, for small ξ, η ,

$$(12) \quad \overline{U} = 4(1 - \mu) + 4(\mu^2 + \mu - \frac{1}{2}C)(\xi^2 + \eta^2) + (\xi, \eta)_4,$$

* The function \overline{U} , defined by (9₂) below, has nothing to do with the function U which is defined by (1₂)–(1₃); the latter U will not be used in what follows.

where $(\xi, \eta)_4$ denotes a regular power series which begins with terms of the fourth order in (ξ, η) and has real coefficients which depend only on μ . In particular, \bar{U} remains regular at the point $(\xi, \eta) = (0, 0)$ into which the singular point $(x, y) = (-\mu, 0)$ of (5₂)-(6₁) is transformed by $x + iy = -\mu + \zeta^2$. This means that, as expected, the isoenergetic transition from (6₁), (6₂) to (10₁), (9₁) eliminates the singularity at the mass $1 - \mu$.

§448. In order to see what happens to the path $x = x(t)$, $y = y(t)$ at the date $t = t_0$ of a collision with the body $1 - \mu$, assign to a fixed value of the time variable \bar{t} of (10₁), say to $\bar{t} = 0$, four initial values $\xi_0, \eta_0, \dot{\xi}_0, \dot{\eta}_0$ in such a way that (ξ_0, η_0) is the position $(0, 0)$ of the body $1 - \mu$, while $(\dot{\xi}_0, \dot{\eta}_0)$ satisfies the energy condition (9₁). This means that

$$(13) \quad \xi_0 = 0, \eta_0 = 0; \dot{\xi}_0 = (8 - 8\mu)^{\frac{1}{2}} \cos \gamma, \dot{\eta}_0 = (8 - 8\mu)^{\frac{1}{2}} \sin \gamma, \\ (0 \leq \mu < 1),$$

holds for a suitable γ , which is, therefore, the only integration constant not disposed of. Actually, the energy (7₅) is another integration constant, since it occurs explicitly in the force function (11) of (10₁).

For reasons of regularity, the coordinates of the collision path $\xi = \xi(\bar{t})$, $\eta = \eta(\bar{t})$ may be developed according to positive powers of \bar{t} into series which are convergent for small $|\bar{t}|$, i.e., for all dates close enough to the date $\bar{t} = 0$ of collision. In view of (13), these Taylor series begin with

$$(14) \quad \begin{aligned} \xi &= ((8 - 8\mu)^{\frac{1}{2}} \cos \gamma) \cdot \bar{t} + \dots, \\ \eta &= ((8 - 8\mu)^{\frac{1}{2}} \sin \gamma) \cdot \bar{t} + \dots; \end{aligned}$$

so that, since $x = -\mu + \xi^2 - \eta^2$, $y = 2\xi\eta$ (cf. §447),

$$(15) \quad \begin{aligned} x &= -\mu + (8(1 - \mu) \cos 2\gamma) \cdot \bar{t}^2 + \dots, \\ y &= (8(1 - \mu) \sin 2\gamma) \cdot \bar{t}^2 + \dots. \end{aligned}$$

Furthermore, $\xi^2 + \eta^2 = 8(1 - \mu)\bar{t}^2 + \dots$, by (14); so that, from (10₂),

$$(16) \quad t = \frac{3}{8}(1 - \mu)\bar{t}^3 + \dots, \quad (0 \leq \mu < 1)$$

if, without loss of generality, $t = 0$ is chosen to belong to $\bar{t} = 0$.

§449. Clearly, (16) has, in the neighborhood of the date $\bar{t} = 0$ of the collision, a unique inverse $\bar{t} = \bar{t}(t)$ which may be developed for small $t \geq 0$ into a real power series in $\sqrt[3]{t} \geq 0$. On substituting this Puiseux expansion of $\bar{t} = \bar{t}(t)$ into (15), one sees that the nature of the singularity of the coordinates $x = x(t)$, $y = y(t)$ at the date $t = 0$ of the collision is the same as in §269 (or §414). In particular, (15)–(16) represents a uniformization of $x = x(t)$, $y = y(t)$ at $t = 0$; so that the motion is, by means of real analytic continuation, uniquely defined for dates t which follow the date $t = 0$ of collision.

§450. Since $1 - \mu > 0$, it is also seen from (15) that the path in the synodical (x, y) -plane acquires a cusp at the date of the collision. By this is meant that the particle reaches the mass $1 - \mu$, which rests at $(x, y) = (-\mu, 0)$, in a definite direction, and is rejected by the mass $1 - \mu$ in the same direction. In fact, this direction is determined by the (arbitrary) integration constant γ of (13).

§451. It is clear from §448–§450 that what is essential in the mapping $x + iy = z(\zeta)$, is not its explicit form $x + iy = -\mu + \zeta^2$, but merely the fact that the singularity $(x, y) = (-\mu, 0)$ of (5₂) is mapped by the inverse of $x + iy = z(\zeta)$ on a point (ξ, η) at which the derivative $z_\zeta(\zeta)$ of the single-valued regular function $z(\zeta) \equiv z(\xi + i\eta)$ vanishes in the first order. (This means that the mapping ceases to be conformal in such a way that the angles are doubled.) Since a similar remark holds for the singularity of (5₂) at $(x, y) = (1 - \mu, 0)$, it follows that, if the mapping function $z = z(\zeta)$ is chosen as in (31), §56, the singularities at both bodies $\mu, 1 - \mu$ will be regularized; so that one can use the same variables $\xi, \eta; \bar{t}$ in case of a collision with either of the mass $\mu, 1 - \mu$. In fact, the derivative z_ζ of (31), §56 vanishes, and then in the first order, if and only if $(\xi, \eta) = (0, 0), (\pm \pi, 0), (\pm 2\pi, 0), \dots$; and these points are mapped by (31), §56 alternately on the two points $(x, y) = (-\mu, 0), (1 - \mu, 0)$.

In order to obtain the explicit form of (8₁)–(8₂) in case of the mapping (31), §56, one has merely to observe that, by (32₂), §56,

$$(17_1) \quad |z_\zeta|^2 = \frac{1}{8}(\cosh 2\eta - \cos 2\xi); \quad (17_2) \quad \dot{t} = |z_\zeta|^2, \text{ by (8}_2\text{)};$$

and that substitution of (17₁) and (5₂) into (9₂) gives

$$\bar{U} = \frac{1}{2}(\cosh \eta - (1 - 2\mu) \cos \xi)$$

$$\begin{aligned}
 (18) \quad & + \frac{1}{16}(1 - 2\mu + \mu^2 - C)(\cosh 2\eta - \cos 2\xi) \\
 & + \frac{1}{256}(\cosh 4\eta - \cos 4\xi) \\
 & + \frac{1}{64}(1 - 2\mu)(\cosh 3\eta \cos \xi - \cosh \eta \cos 3\xi),
 \end{aligned}$$

in view of (30)–(34), §56 and of $2 \cos^2 \alpha \equiv 1 + \cos 2\alpha$, $4 \cos^3 \alpha \equiv 3 \cos \alpha + \cos 3\alpha$.

Substitution of (17₁) and (18) into (8₁), (9₁) supplies the explicit form of the equations of motion for every fixed C . Notice that (17₁) and (18) are regular analytic in the whole (ξ, η) -plane.

§452. In the case $\mu = \frac{1}{2}$ of two equal masses, (18) simplifies to

$$\begin{aligned}
 (18 \text{ bis}) \quad \bar{U} &= \frac{1}{2} \cosh \eta + \frac{1}{64}(1 - 4C)(\cosh 2\eta - \cos 2\xi) \\
 &+ \frac{1}{256}(\cosh 4\eta - \cos 4\xi).
 \end{aligned}$$

The numerical calculations carried out at the Copenhagen Observatory, which deal with this symmetric case $\mu = 1 - \mu$, are based on the equations (8₁), (9₁) belonging to (18 bis) and (17₁).

§453. It is clear from the beginning of §451 that the mapping (25), §55 can be used for the same purpose as (31), §56. The representation of \bar{U} and $|z_\zeta|^2$ in the case (25), §55 has over (18) and (17₁) the advantage of leading to algebraic, instead of to transcendental, functions. The correspondence between (x, y) and (ξ, η) is now one-to-two (instead of being, as in §451, one-to-infinity) and can, therefore, conveniently be used in topological discussions in the large (cf. §500 below).

§454. In view of the beginning of §451, it is natural to ask, what would happen if one replaced the mapping $z = -\mu + \zeta^2$ of §447 by $z = -\mu + \zeta^n$, where n is an integer exceeding 2. The answer is that this mapping is useless for the purpose of regularization.

In fact, if $n > 2$, one readily sees from the deduction of (12) that \bar{U} , instead of having, as there, a constant term ($= 4 - 4\mu \neq 0$), vanishes at $(\xi, \eta) = (0, 0)$, i.e., at the point at which the collision takes place. Hence, (9₁) requires that (13) be replaced by $\xi_0 = 0$, $\eta_0 = 0$; $\dot{\xi}_0 = 0$, $\dot{\eta}_0 = 0$. But $\xi(\bar{t}) \equiv 0$, $\eta(\bar{t}) \equiv 0$ is one, hence the only, solution of (8₁) which satisfies this initial condition; in fact, \bar{U}_ξ , \bar{U}_η are readily seen to vanish at $(\xi, \eta) = (0, 0)$, not only in the case (12) of $n = 2$ but for any $n \geq 2$.

Accordingly, if $n > 2$, the singular point at which the mass $1 - \mu$ rests is transformed into an equilibrium solution with reference to

the time variable \bar{t} . Hence, the collision which takes place at a finite t -date, say at $t = 0$, will not take place at a finite \bar{t} -date but as $\bar{t} \rightarrow \infty$, that is, asymptotically (cf. the end of §167). In other words, the denominator of (8₂), when considered as a function of t , vanishes at the date $t = 0$ of the collision too strongly, if $n > 2$.

§455. Returning to (5₁)–(6₂), suppose that the constants μ , C have fixed values, and that the position (x, y) of the third particle is varying in such a way that the value of U remains bounded. In view of (6₂), this will be the case if and only if x' and y' remain bounded. On the other hand, (5₂) shows that U_x and U_y remain bounded if and only if so do both distances r_i and their reciprocal values $1/r_i$, where

$$(19) \quad r_1 = \{(x + \mu)^2 + y^2\}^{\frac{1}{2}}, \quad r_2 = \{(x + \mu - 1)^2 + y^2\}^{\frac{1}{2}};$$

(cf. §443 bis).

Hence, on writing (6₁) as a system of four differential equations of the first order for x, y, x', y' , and placing, along a fixed solution $x = x(t), y = y(t)$ of (6₁),

$$(20) \quad \rho(t) = \text{Min } (r_1(t), r_2(t), 1/r_1(t), 1/r_2(t)),$$

one readily arrives at the following analogue to the lemma formulated at the end of §408:

If the value of μ and of the integration constant C in (6₂) are fixed, there exist for every positive number ρ^* two positive numbers α^*, β^* such that any solution $x = x(t), y = y(t)$ of (6₁)–(6₂) for which the inequality $\rho(t) > \rho^*$ is satisfied at some $t = t_0$ is a solution which not only exists and is regular analytic for every t contained in the interval $|t - t_0| < \alpha^*$, but is, in addition, such as to satisfy the inequalities

$$(21_1) \quad (x(t) - x(t_0))^2 + (y(t) - y(t_0))^2 < \beta^*; \quad (21_2) \quad \rho(t) > \frac{1}{2}\rho^*$$

for every t between $t_0 - \alpha^*$ and $t_0 + \alpha^*$. As in §408, the point is that α^*, β^* do not depend on the choice of t_0 (but merely on μ, C and ρ^*).

§456. Suppose that a solution $x = x(t), y = y(t)$ of the restricted problem of three bodies ceases either to exist or to be regular analytic (in t), when t tends, say decreasingly, to a fixed finite $t = t^0$, say to $t^0 = 0$. Then, as a consequence of the lemma expressed by (21₁)–

(21₂), one must have $\lim_{t \rightarrow +0} \rho(t) = 0$, as $t \rightarrow +0$. The proof is the same as in §409. Actually, not only $\lim_{t \rightarrow +0} \rho(t) = 0$ holds but also $\lim_{t \rightarrow +0} \rho(t) = 0$, as $t \rightarrow +0$. The proof is the same as in §409. But comparison of (19) with (20) shows that $\lim_{t \rightarrow +0} \rho(t) = 0$ holds if and only if one of the three conditions $\lim_{t \rightarrow +0} r_1(t) = 0$, $\lim_{t \rightarrow +0} r_2(t) = 0$, $\lim_{t \rightarrow +0} r_1(t) = +\infty$, which are mutually exclusive, is satisfied. In the first and the second cases, one has to do with a collision with the masses $1 - \mu$ and μ , respectively. These two cases, which are equivalent, have been treated in §447–§449. And it will now be shown that this case of an ordinary collision of the moving particle with one of the two resting bodies $1 - \mu$, μ exhausts all possibilities, i.e., that the third case, that in which $\lim_{t \rightarrow +0} r_1(t) = +\infty$, can never occur.

§457. In order to prove this, suppose, if possible, that $r_1(t) \rightarrow +\infty$, as $t \rightarrow +0$. This assumption is, by (19), equivalent to $r_2(t) \rightarrow +\infty$, and also to $x(t)^2 + y(t)^2 \rightarrow +\infty$, when $t \rightarrow +0$. Hence, (5₂) shows that if $t (> 0)$ is close to $t = 0$, the contribution of the gravitational terms of U to the force vector (U_x, U_y) is very small; while the principal part of (U_x, U_y) is the large force vector (x, y) which represents the gradient of the centrifugal term $\frac{1}{2}(x^2 + y^2)$ of (5₂). This means that, as $t \rightarrow +0$, a close approximation to (6₁)–(6₂) is represented by

$$(22_1) \quad \bar{x}'' - 2\bar{y}' - \bar{x} = 0, \quad \bar{y}'' + 2\bar{x}' - \bar{y} = 0;$$

$$(22_2) \quad \bar{x}'^2 + \bar{y}'^2 = \bar{x}^2 + \bar{y}^2 - C.$$

But (22₁) is a homogeneous linear system with constant coefficients. Hence, every solution $\bar{x} = \bar{x}(t)$, $\bar{y} = \bar{y}(t)$ of (22₁) is regular analytic for every finite t ; and so $\bar{x}(t)^2 + \bar{y}(t)^2$ must tend to a finite limit, as $t \rightarrow +0$. On the other hand, on estimating the deviation of (6₁) from (22₁) on the assumption that $x(t)^2 + y(t)^2 \rightarrow +\infty$ as $t \rightarrow +0$, one readily sees from the appraisals which are supplied by adapting the standard procedure of successive approximations, that also $\bar{x}(t)^2 + \bar{y}(t)^2 \rightarrow +\infty$ as $t \rightarrow +0$. This contradiction proves that, as stated at the end of §456, one cannot have $x(t)^2 + y(t)^2 \rightarrow +\infty$ as $t \rightarrow +0$.

§458. On comparing §457 with §456, and using (20)–(22₂) again, one sees that not only $\lim_{t \rightarrow +0} \{x(t)^2 + y(t)^2\} = +\infty$ but also $\overline{\lim}_{t \rightarrow +0} \{x(t)^2 + y(t)^2\} = +\infty$ is impossible, when t tends to a finite t_0 , say to $t_0 = 0$. In other words, the coordinates $x(t)$, $y(t)$ must remain

bounded for every solution of the restricted problem of three bodies, so long as t varies over a finite range.†

§459. The considerations of §456–§458 tacitly assume that the interior of the t -range under consideration is free of dates of collision. Actually, everything remains valid also when this assumption is not made. The proof proceeds as follows:

According to §449, there is a unique analytic continuation of the motion through any date of collision. On comparing this fact with the one mentioned at the end of §456, one sees that the motion can always be defined for $-\infty < t < +\infty$, if the dates of collision do not cluster at a finite $t = t^*$. And it will be shown that such a finite t^* can never exist; i.e., that the dates of collisions, if they exist at all, form either a finite sequence of points on the t -axis or an infinite sequence which tends to $\pm\infty$ (possibly only to $+\infty$ or only to $-\infty$).

§460. Suppose, if possible, that, for a given solution $x = x(t)$, $y = y(t)$ of (6₁), there happen infinitely many collisions on a finite t -interval which has the cluster value $t^* \neq \pm\infty$ of the dates of collisions as an end point. Denoting the successive dates of collisions by t_1, t_2, \dots , one can assume that $t_n > t_{n+1}$ for $n = 1, 2, \dots$, and that $t_n \rightarrow 0$ as $n \rightarrow \infty$; so that $t^* = 0$, while there is no collision between $t = t_n$ and $t = t_{n+1}$.

According to (19), either $r_1(t)$ or $r_2(t)$ vanishes at every $t = t_n$; so that $\rho(t_n) = 0$, by (20). Since $t_n \rightarrow +0$ as $n \rightarrow \infty$, it follows, by letting $n \rightarrow \infty$ in $\rho(t_n) = 0$, that $\underline{\lim} \rho(t) = 0$ as t tends to $+0$ continuously. But if one proceeds in the same way as at the beginning of §456, one sees that $\underline{\lim} \rho(t) = 0$ again implies that $\lim \rho(t) = 0$, as t tends to $+0$ continuously.

Hence, a repetition of the considerations of §456–§458 shows that, as t tends to $+0$ continuously, either $\lim r_1(t) = 0$ or $\lim r_2(t) = 0$. For reasons of symmetry, it is sufficient to consider the first of these two cases (which are, by (19), mutually exclusive). But if $\lim r_1(t) = 0$ as $t \rightarrow +0$, the regularization (10₁)–(12) of the problem is readily applicable at $t = 0$. Hence, §449 shows that $r_1 = r_1(t)$ has at $t = 0$ an algebraic singularity. Consequently, the function $r_1(t)$ cannot attain the value 0 at dates t which cluster at $t = 0$. But this contradicts the assumption that $\rho(t_n) = 0$ for infinitely many t_n .

† That this is not evident itself, is seen from the footnote to §186.

which cluster at $t = 0$. This contradiction completes the proof of the statement made at the end of §459.

§461. The above results may be summarized as follows:

Every solution $x = x(t)$, $y = y(t)$ of the restricted problem of three bodies exists for $-\infty < t < +\infty$, the real finite singularities being necessarily collisions of the third body with one of the two bodies $1 - \mu$, μ . In fact, the motion admits, by §449, of a unique real analytic continuation through a date of collision (if any); and if there are infinitely many dates t_n of subsequent collisions, then $|t_n| \rightarrow \infty$ as $n \rightarrow \infty$, by §460.

It follows that the regularization of an arbitrary solution $x = x(t)$, $y = y(t)$ of the restricted problem of three bodies in terms of the variables of §451 or of §453 is valid for $-\infty < t < +\infty$. Since the t_n cannot have a finite cluster value, it is also seen that the regularizing time variable $\bar{t} = \bar{t}(t)$, which is defined by (8₂) up to an additive constant, runs with t from $-\infty$ to $+\infty$, whether the mapping $z = z(\bar{t})$ is that of §451 or of §453.

The Syzygical Potential Curve

§462. The object of the following considerations (up to §473) is the study of the field of force which is generated by the centrifugal and gravitational forces together. This field of force is the 2-vector function whose components are $U_x(x, y)$, $U_y(x, y)$, where, by (5₂), §443,

$$(1_1) \quad U(x, y) = \frac{1}{2}(x^2 + y^2) + (1 - \mu)\rho^{-1} + \mu\sigma^{-1};$$

$$(1_2) \quad \rho = |(x + \mu)^2 + y^2|^{\frac{1}{2}}, \quad \sigma = |(x + \mu - 1)^2 + y^2|^{\frac{1}{2}}.$$

It is convenient to visualize $U = U(x, y)$ as a surface, situated over the (x, y) -plane of an (x, y, U) -space. By (1₁)-(1₂), there exists a different surface $U = U(x, y)$ for every μ . The limiting case of §300 will be excluded; so that $0 < \mu < 1$.

It is clear from (1₁)-(1₂) that the ordinate U of the surface is everywhere positive, and becomes $+\infty$ only at the points occupied by the two masses

$$(2_1) \quad 1 - \mu: (x, y) = (-\mu, 0); \quad (2_2) \quad \mu: (x, y) = (1 - \mu, 0),$$

but tends to $+\infty$ also when $x^2 + y^2 \rightarrow +\infty$. Furthermore, the surface is symmetric with respect to the plane $y = 0$, since $U(x, y) = U(x, -y)$, by (1₁)-(1₂).

This plane cuts the potential surface $U = U(x, y)$ along a curve $U = U(x, 0)$, plotted against the axis x of syzygies. Up to §468, only this syzygical potential curve will be considered; although it will be thought of as embedded into the surface, since also the function $U_{yy}(x; 0)$ of x will be studied.

§463. It is easily verified from (1₁)–(1₂) that

$$(3_1) \quad U(x, 0) = \frac{1}{2}x^2 + \frac{1 - \mu}{|x + \mu|} + \frac{\mu}{|x + \mu - 1|};$$

$$(3_2) \quad U_x(x, 0) = x - (1 - \mu) \frac{x + \mu}{|x + \mu|^3} - \mu \frac{x + \mu - 1}{|x + \mu - 1|^3};$$

and that, since (1₂) reduces to $\rho = |x + \mu|$, $\sigma = |x + \mu - 1|$,

$$(4_1) \quad U_{xx}(x, 0) = 1 + \frac{1 - \mu}{\frac{1}{2}\rho^3} + \frac{\mu}{\frac{1}{2}\sigma^3};$$

$$(4_2) \quad U_{yy}(x, 0) = 1 - \frac{1 - \mu}{\rho^3} - \frac{\mu}{\sigma^3};$$

$$(4_3) \quad U_y(x, 0) \equiv 0 \equiv U_{xy}(x, 0);$$

(4₃) being obvious from $U(x, y) = U(x, -y)$.

Notice that the two points (2₁)–(2₂) subdivide the x -axis into the three regions

$$(5_I) \quad -\infty < x < -\mu;$$

$$(5_{II}) \quad -\mu < x < 1 - \mu;$$

$$(5_{III}) \quad 1 - \mu < x < +\infty$$

within which one respectively has

$$(6_I) \quad \rho = -(\mu + x), \quad \sigma - \rho = 1;$$

$$(6_{II}) \quad \rho = \mu + x, \quad \sigma + \rho = 1;$$

$$(6_{III}) \quad \rho = \mu + x, \quad \rho - \sigma = 1.$$

Correspondingly, (3₂) may be written in the regions (5_I), (5_{II}) as

$$(7_I) \quad U_x(x, 0) = -\mu - \rho + (1 - \mu)/\rho^2 + \mu/(1 + \rho)^2;$$

$$(7_{II}) \quad U_x(x, 0) = -\mu + \rho - (1 - \mu)/\rho^2 + \mu/(1 - \rho)^2;$$

while $U_x(x, 0)$ in (5_{III}) follows by writing

$$(7_{\text{III}}) \quad \mu, \sigma, \rho; -U_x \text{ instead of } 1 - \mu, \rho, \sigma; U_x \text{ in } (7_{\text{I}}).$$

For this reason of symmetry, it will always be sufficient to consider the first two, instead of all three, regions (5_I)–(5_{III}) of the axis of syzygies.

§464. It will now be shown that, for every fixed value of the positive mass parameter μ (< 1), the function (3₂) of x has in each of the three regions (5_k) exactly one zero, $x = x_k$, which is, of course, a function $x_k(\mu)$ of μ . Furthermore, it will be shown that

$$(8_{\text{I}}) \quad -1 - \mu < x_{\text{I}}(\mu) < -\mu;$$

$$(8_{\text{II}}) \quad -\mu < x_{\text{II}}(\mu) < 1 - \mu;$$

$$(8_{\text{III}}) \quad 1 - \mu < x_{\text{III}}(\mu) < 2 - \mu;$$

so that the distance between any of the three points $(x, y) = (x_k(\mu), 0)$ of the x -axis and at least one of the two masses (2₁)–(2₂) is less than the distance, 1 ($= |(1 - \mu) - (-\mu)|$), between the two masses (2₁)–(2₂). In other words, all three $\min(\rho_k, \sigma_k) < 1$ for every μ , where it is understood that $\rho_k = \rho_k(\mu)$, $\sigma_k = \sigma_k(\mu)$ are defined as the values of the distances (1₂) for $(x, y) = (x_k(\mu), 0)$; $k = \text{I, II, III}$.

First, $0 < \mu < 1$; so that (4₁) is positive for every x . Since (4₁) is the derivative of $U_x(x, 0)$, it follows that the function (3₂) of x is increasing at every x . However, (3₂) becomes infinite at the two points (2₁), (2₂). Since these separate the x -axis into the three regions (5_k), it follows that $U_x(x, 0)$ is a steadily increasing continuous function on each of the three intervals (5_k). But $U_x(x, 0)$ tends to $-\infty$ or to $+\infty$ according as x tends to the lower or the upper end of any of these three intervals; in fact, (3₂) shows that $U_x(\pm\infty, 0) = \pm\infty$, $U_x(-\mu \pm 0, 0) = \mp\infty$, $U_x(1 - \mu \pm 0, 0) = \mp\infty$. Consequently, $U_x(x, 0)$ attains on each of the three intervals (5_k) every value between $-\infty$ and $+\infty$, hence also the value 0, exactly once. This proves the existence and uniqueness of the three $x_k = x_k(\mu)$.

It is clear from this proof that if x is any point of (5_k), then $U_x(x, 0) \lessgtr 0$ according as $x \lessgtr x_k(\mu)$. This implies that the point $x = x_k(\mu)$ subdivides the region (5_k) into two subintervals in such a way that the potential function $U(x, 0)$ is steadily decreasing on the first, and steadily increasing on the second, of these subintervals. In other words, the positive function (3₁), which becomes $+\infty$ at

both ends of (5_k) , has at $x = x_k(\mu)$ a minimum and is convex (from below) on (5_k) .

It follows that in order to prove that (8_I) is satisfied by the point $x = x_I(\mu)$ of (5_I) , it is sufficient to show that $U_x(x, 0) < 0$ at the end $x = -1 - \mu$ of (8_I) . But this condition is satisfied, since $U_x(-1 - \mu, 0) = -\frac{7}{4}\mu$, by (3_2) . This proves (8_I) ; and (8_{III}) is, by the end of §463, equivalent to (8_I) . Finally, (8_{II}) does not improve on the fact that $x = x_{II}(\mu)$ lies in (5_{II}) . This completes the proof of the three inequalities $\min(\rho_k, \sigma_k) < 1$, which are equivalent to (8_k) ; $k = I, II, III$.

§464 bis. As a consequence, one has, for every k and μ ,

$$(9_1) \quad U_{yy}(x_k(\mu), 0) < 0; \quad (9_2) \quad (1 - \mu)/\rho_k^3 + \mu/\sigma_k^3 > 1.$$

First, (9_1) is, by (4_2) , equivalent to (9_2) . Next, if $k = I$, then $\sigma_I = 1 + \rho_I$, by (6_I) . Hence, if $k = I$, the sum on the left of (9_2) is greater than $(1 - \mu)/\rho_I^3 + \mu/\rho_I^3$; and so greater than 1, if $\rho_I < 1$. But $\rho_I < 1$ is implied by the end of §464. This proves (9_2) for $k = I$, hence, by the end of §463, for $k = III$ also. Finally, if $k = II$, then $\rho_k + \sigma_k = 1$, by (6_{II}) ; so that both positive numbers ρ_k, σ_k are less than 1, and therefore (9_2) is obvious.

§465. It will now be shown that all three $\rho_k(\mu)$ and all three $\sigma_k(\mu)$ are strictly monotone functions of μ on the whole range $0 < \mu < 1$.

By the end of §463, it is sufficient to prove this for $k = I, II$; so that $\sigma_k(\mu) = 1 \pm \rho_k(\mu)$, by (6_I) – (6_{II}) . Hence, it is sufficient to prove that $\rho_k = \rho_k(\mu)$ has a finite non-vanishing derivative $d\rho_k/d\mu$ for $k = I, II$ and $0 < \mu < 1$.

To this end, notice first that, by (7_I) – (7_{II}) and the definition of the ρ_k ,

$$(10) \quad 0 = -\mu \mp \rho_k \pm (1 - \mu)/\rho_k^2 + \mu/(1 \pm \rho_k)^2; \quad (k = I, II),$$

where the upper signs belong to $k = I$ and the lower to $k = II$. Differentiating the identity (10) in μ , where $\rho_k = \rho_k(\mu)$, with respect to μ , one obtains

$$(11) \quad \mp 1 - \frac{1}{\rho_k^2} \pm \frac{1}{(1 \pm \rho_k)^2} = \left\{ 1 + \frac{2(1 - \mu)}{\rho_k^3} + \frac{2\mu}{(1 \pm \rho_k)^3} \right\} \frac{d\rho_k}{d\mu};$$

($k = I, II$).

But $1 \pm \rho_k = \sigma_k > 0$, by (6_I)–(6_{II}); so that the coefficient $\left\{ \right\}$ of $d\rho_k/d\mu$ on the right of (11) is positive. Hence, in order to infer from (11) that $\rho_k = \rho_k(\mu)$ has a finite non-vanishing derivative with respect to μ , it is sufficient to show that the expression on the left of (11) cannot vanish. Since $1 \pm \rho_k = \sigma_k$ for $k = I, II$, respectively, it follows that it is sufficient to prove the inequalities $1 - 1/\sigma_I^2 \neq -1/\rho_I^2$ and $1 - 1/\rho_{II}^2 \neq 1/\sigma_{II}^2$. Hence, one need prove only that each of the three positive numbers $1/\sigma_I$; ρ_{II} , σ_{II} is less than 1. But $1 < 1 + \rho_I = \sigma_I$ and $\rho_{II} + \sigma_{II} = 1$, by (6_I)–(6_{II}); so that the proof is complete.

§465 bis. The result of §465 may be completed by calculating the limiting values of the six monotone functions $\rho_k(\mu)$, $\sigma_k(\mu)$ at the end points of the interval $0 < \mu < 1$. These limiting values are

$$(12_I) \quad \rho_I(+0) = 1, \quad \rho_I(1-0) = 0;$$

$$(12_{II}) \quad \rho_{II}(+0) = 1, \quad \rho_{II}(1-0) = 0;$$

$$(12_{III}) \quad \sigma_{III}(+0) = 0, \quad \sigma_{III}(1-0) = 1.$$

$$(12_I^*) \quad \sigma_I(+0) = 2, \quad \sigma_I(1-0) = 1;$$

$$(12_{II}^*) \quad \sigma_{II}(+0) = 0, \quad \sigma_{II}(1-0) = 1;$$

$$(12_{III}^*) \quad \rho_{III}(+0) = 1, \quad \rho_{III}(1-0) = 2.$$

In fact, (12_I)–(12_{II}) follow from (10), where $k = I, II$, by letting $\mu \rightarrow +0$ and $\mu \rightarrow 1-0$. And (12_{III}) is, by the end of §463, implied by (12_I). Finally, (12_k^{*}) is, by (6_k), equivalent to (12_k), where $k = I, II, III$.

§466. Next, the relative magnitude of the values of the functions $\rho_k(\mu)$, $\sigma_k(\mu)$ for any fixed μ will be determined. It will be shown that, while

$$(13_1) \quad \sigma_{III}(\mu) \leq \rho_I(\mu) \text{ for } \mu \leq \frac{1}{2}; \quad (13_2) \quad \sigma_{II}(\mu) \leq \rho_{II}(\mu) \text{ for } \mu \leq \frac{1}{2},$$

one has

$$(14) \quad \rho_{II}(\mu) < \rho_I(\mu) \text{ for } 0 < \mu < 1.$$

(It is understood that, for reasons of symmetry, the relations (13₁), (14) imply equivalent relations.)

First, $\mu = \frac{1}{2}$ means that the two masses μ , $1 - \mu$ are equal. Hence,

(13₁) and (13₂) follow from the definitions (§464) for reasons of symmetry. In fact, (12_{II}^{*}), (12_{III}) and (12_I), (12_{II}) imply that the functions $\sigma_{II}(\mu)$, $\sigma_{III}(\mu)$ and $\rho_I(\mu)$, $\rho_{II}(\mu)$, which are strictly monotone by §465, are increasing and decreasing, respectively.

In order to prove (14), notice first that application of (10) at $\mu = \frac{1}{2}$ shows that $\rho_{II}(\frac{1}{2}) = \frac{1}{2}$, and that $\rho_I(\frac{1}{2})$ is a (positive) root λ of the quintic equation $\phi(\lambda) = 0$, where

$$\phi(\lambda) = 2\lambda^5 + 5\lambda^4 + 4\lambda^3 - \lambda^2 - 2\lambda - 1; \text{ so that } \phi(\frac{2}{3}) < 0 < \phi(1),$$

and so $\phi(\lambda) = 0$ has a root λ between $\frac{2}{3}$ and 1. And this root must be the root $\rho_I(\frac{1}{2})$, since the coefficients of $\phi(\lambda)$ have only one change of sign and are, therefore, incompatible with the existence of more than one positive root. Thus, it is clear from $\rho_{II}(\frac{1}{2}) = \frac{1}{2}$ that (14) is true at $\mu = \frac{1}{2}$. Hence, (14) is true for every μ between 0 and 1, unless $\rho_I(\mu) = \rho_{II}(\mu)$ at a certain μ , say at $\mu = \mu^*$. But then addition of the two equations (10) shows that the common value ρ^* of $\rho_I(\mu^*)$ and $\rho_{II}(\mu^*)$ must satisfy the condition

$$0 = -2\mu^* + \mu^*/(1 - \rho^*)^2 + \mu^*/(1 + \rho^*)^2;$$

i.e.,

$$\mu^* \cdot (1 - \rho^{*2})^2 = \mu^* \cdot (1 + \rho^{*2}).$$

And this is a contradiction, since $\mu^* > 0$, $\rho^* > 0$.

§467. According to §464, the minimum of $U(x, 0)$ on the interval (5_k) is attained only at the point $x = x_k(\mu)$. It will now be shown that the greatest of the three relative minima always belongs to $k = II$; in fact,

$$(15_1) \quad U(x_l(\mu), 0) < U(x_{II}(\mu), 0) \text{ for } 0 < \mu < 1, \text{ where } l = I, III;$$

$$(15_2) \quad U(x_I(\mu), 0) \leq U(x_{III}(\mu), 0) \text{ according as } \mu \leq \frac{1}{2}.$$

For a fixed μ , let θ denote any value between 0 and $\rho_I = \rho_I(\mu)$ ("between" excluding equality). Then $-\theta - \mu$ lies between $-\rho_I(\mu)$ and $-\mu$. Hence, $-\theta - \mu$ is a point of the interval (5_I) but is, by (6_I), not the point $x_I(\mu)$. Since the minimum of $U(x, 0)$ on the interval (5_I) is attained only at $x_I(\mu)$, it follows that $U(x_I(\mu), 0) < U(-\theta - \mu, 0)$.

On the other hand, the assumption $0 < \theta < \rho_I(\mu)$ also implies that $0 < \theta < 1$, since, as pointed out in §464 bis, one has $\rho_I(\mu) < 1$ for every μ . But it is easily verified from (3₁) that if ϑ is any value be-

tween 0 and 1, then the difference $U(-\vartheta - \mu, 0) - U(\vartheta - \mu, 0)$ is identical with the product $-(1 + \vartheta^2)(1 - \vartheta^2)^{-1}\vartheta\mu$ and is, therefore, negative. Hence, $U(-\theta - \mu, 0) < U(\theta - \mu, 0)$.

On comparing the inequalities found at the ends of the two preceding paragraphs, one sees that $U(x_I(\mu), 0) < U(\theta - \mu, 0)$ holds for any number θ which lies between 0 and $\rho_I(\mu)$. It follows that in order to prove (15_I) for $l = I$, it is sufficient to assure that $x = x_{II}(\mu)$ lies between 0 and $\rho_I(\mu)$. But this is assured, in view of (6_{II}), by (14). This proves (15_I) for $l = I$ and so, for reasons of symmetry (cf. the end of §463), for $l = III$ also.

§467 bis. It will now be shown that the function $U(x_{III}(\mu), 0)$ of μ is steadily decreasing for $0 < \mu < 1$. This will imply (15₂) for reasons of symmetry, since it is then clear, again for reasons of symmetry, that the function $U(x_I(\mu), 0)$ of μ is steadily increasing for $0 < \mu < 1$.

Thus, it is sufficient to show that the total derivative $dU/d\mu$ of $U(x_{III}(\mu), 0)$ is nowhere positive. But this total derivative is identical with the value of the partial derivative $U_\mu(x, 0)$ at $x = x_{III}(\mu)$, since $U_x(x, 0)$ vanishes at $x = x_{III}(\mu)$, by the definition (§464) of $x_{III}(\mu)$. Thus, it is sufficient to prove that $U_\mu(x_{III}(\mu), 0) < 0$ for $0 < \mu < 1$.

To this end, let x be any point of the region (5_{III}). Then, by (3_I),

$$U = \frac{1}{2}x^2 + \frac{1 - \mu}{x + \mu} - \frac{\mu}{x + \mu - 1};$$

hence

$$U_\mu - U_x = -x - \frac{1}{x + \mu} - \frac{1}{x + \mu - 1},$$

as seen by calculating the partial derivatives U_μ, U_x of U . Since (5_{III}) implies that $0 < x$ and $0 < x + \mu - 1 < x + \mu$, it follows that $U_\mu - U_x < 0$ at every point x of (5_{III}). This completes the proof, since $U_x = 0$ at the point $x = x_{III}(\mu)$ of (5_{III}).

§468. In view of (6_k), any two of the three functions $x_k(\mu); \rho_k(\mu), \sigma_k(\mu)$ of μ determine the third for every fixed k . And (10), together with (7_{III}), shows that each of the three functions $\rho_k(\mu)$ of μ is determined by a quintic equation (whose coefficients are linear in μ). The values of the $x_k(\mu)$ in the table were calculated from these quintic

equations, and then the corresponding values of $U = U(x, 0)$, $U_{xx}(x, 0)$, $U_{yy}(x, 0)$ from (3₁), (4₁), (4₂).

μ	x_I	$U_{xx}(x_I, 0)$	$U_{yy}(x_I, 0)$	x_{II}	$U_{xx}(x_{II}, 0)$	$U_{yy}(x_{II}, 0)$	x_{III}	$U_{xx}(x_{III}, 0)$	$U_{yy}(x_{III}, 0)$
0.01	-1.0042	3.0174	-0.0087	0.8481	11.1334	-4.0667	1.1468	7.4670	-2.2335
0.02	-1.0083	3.0356	-0.0178	0.8035	11.7846	-4.3923	1.1801	7.1264	-2.0632
0.03	-1.0125	3.0532	-0.0266	0.7696	12.2500	-4.6250	1.2012	6.8944	-1.9472
0.04	-1.0167	3.0710	-0.0355	0.7409	12.6380	-4.8190	1.2164	6.7142	-1.8571
0.05	-1.0208	3.0898	-0.0449	0.7152	12.9658	-4.8929	1.2281	6.5594	-1.7797
0.10	-1.0416	3.1834	-0.0917	0.6090	14.1750	-5.5875	1.2597	6.0134	-1.5067
0.20	-1.0828	3.3856	-0.1928	0.4381	15.5972	-6.2986	1.2710	5.3308	-1.1654
0.30	-1.1232	3.6086	-0.3043	0.2861	16.4154	-6.7077	1.2567	4.8488	-0.9244
0.40	-1.1620	3.8584	-0.4292	0.1416	16.8588	-6.9294	1.2308	4.4640	-0.7320
0.50	-1.1984	4.1396	-0.5698	0.0000	17.0000	-7.0000	1.1984	4.1434	-0.5717

The Potential Surface

§469. In §463–§468 the symmetric surface $U = U(x, y)$ of §462 was studied, for every fixed μ , along its plane $y = 0$ of symmetry. In particular, it was shown in §464 that the intersection $U = U(x, 0)$ of this surface and of the plane $y = 0$ is a curve which is convex (from below) on each of the three regions (5_k) of the x -axis; and that the function $U(x, 0)$, which becomes $+\infty$ at both ends of (5_k), attains its minimum on (5_k) at the point $x = x_k(\mu)$ of (5_k). The relative magnitude of these three minima is described by (15₁)–(15₂).

It will now be shown that, for every fixed value of the parameter μ in (1₁)–(1₂), where $0 < \mu < 1$,

(i) there exist in the (x, y) -plane exactly five points at which the tangent plane of the surface $U = U(x, y)$ is parallel to the (x, y) -plane;

(ii) these 5 points (x, y) are the 3 + 2 points

$$(16_1) \quad (x, y) = (x_k(\mu), 0); \quad k = I, II, III;$$

$$(16_2) \quad (x, y) = (\tfrac{1}{2} - \mu, \pm \tfrac{1}{2}\sqrt{3}),$$

(16₁) representing the 3 points considered in §464, and (16₂) those 2 points either of which forms with the two masses (2₁)–(2₂) an equilateral triangle;

(iii) the Hessian matrix of the function $U(x, y)$ at the 5 points (x, y) which satisfy $U_x = 0 = U_y$ is, respectively,

$$(17_1) \quad \begin{pmatrix} U_{xx} & U_{xy} \\ U_{xy} & U_{yy} \end{pmatrix} = \begin{pmatrix} + & 0 \\ 0 & - \end{pmatrix} \text{ at } (x, y) = (x_k(\mu), 0); \quad k = I, II, III$$

$$(17_2) \begin{pmatrix} U_{xx} & U_{xy} \\ U_{xy} & U_{yy} \end{pmatrix} = \frac{\sqrt{27}}{4} \begin{pmatrix} 1/\sqrt{3} & 1 - 2\mu \\ 1 - 2\mu & \sqrt{3} \end{pmatrix} \text{ at } (x, y) = (\tfrac{1}{2} - \mu, \pm \tfrac{1}{2} \sqrt{3}),$$

+ in (17₁) denoting a positive, and - a negative, function of μ and k ;

(iv) the surface $U = U(x, y)$ has* at any of the three points (16₁) a saddle point (and so no relative extremum), while it has a relative minimum at the points (16₂);

(v) while the function $U(x, y) \equiv U(x, -y)$ defined by (1₁)–(1₂) becomes $+\infty$ at both points (2₁)–(2₂) and as $x^2 + y^2 \rightarrow +\infty$, the absolute minimum of $U(x, y)$ in the whole (x, y) -plane is attained at the points (16₂) and has the value $\frac{1}{2}(3 - \mu + \mu^2)$.

§469 bis. In order to prove (i)–(v), notice first that differentiation of (1₁)–(1₂) with respect to x and y gives

$$(18_1) \quad U_x = xV + (1 - \mu)\mu \left(\frac{1}{\sigma^3} - \frac{1}{\rho^3} \right); \quad (18_2) \quad U_y = yV;$$

$$(18_3) \quad V = 1 - \frac{1 - \mu}{\rho^3} - \frac{\mu}{\sigma^3}.$$

The statement (i) deals with the points (x, y) at which the functions (18₁), (18₂) vanish simultaneously. And (18₂) vanishes if and only if either $y = 0$ or $V = 0$. In the first case, where $y = 0$, the vanishing of (18₁) means that x satisfies the condition $U_x(x, 0) = 0$. This, when compared with the definition of $x_k(\mu)$ in §464, supplies the three points (16₁). In the second case, where $V = 0$, one sees that (18₁) vanishes if and only if $\rho = \sigma$; while $V = 0$ and $\rho = \sigma$ imply, by (18₃), that $\rho = \sigma = 1$. Since $\rho = \sigma = 1$ is, in view of (1₂), equivalent to (16₂), the proof of (i)–(ii) is complete.

Next, (17₁) is clear from (4₁), (4₃), (9₁); while (17₂) follows by substituting (16₂) into the second derivatives of U ; cf. (1₁)–(1₂). This proves (iii). And (iv) follows by observing that the two characteristic numbers of the matrix (17₁) are of opposite sign, while those of (17₂) are positive, having the values

$$(18 \text{ bis}) \quad \frac{3}{2}(1 \pm \sqrt{1 - 3\mu(1 - \mu)}),$$

where $\mu(1 - \mu) \leq \frac{1}{4}$, since $0 < \mu < 1$.

* In view of the index relations of Birkhoff-Morse concerning critical points, the facts collected under (iv)–(v) are not quite independent of each other.

Finally, (v) is clear from (i) and (iv); the value of U at the points (16₂) being $\frac{1}{2}\{3 - \mu(1 - \mu)\} (> 0)$, by (1₁)–(1₂).

§470. In case of two equal masses $\mu, 1 - \mu$, the surface $U = U(x, y)$ has, besides the plane of symmetry $y = 0$, the plane of symmetry $x = 0$. In fact, it is seen from (1₁)–(1₂) that if $\mu = \frac{1}{2}$, then not only $U(x, -y) = U(x, y)$ but also $U(-x, y) = U(x, y)$.

In §462–§469, the two equivalent limiting cases $\mu = 0$; $\mu = 1$ of two positive masses $\mu, 1 - \mu$ have been excluded. If $\mu = 0$, then (1₁)–(1₂) reduce to $U = \frac{1}{2}\rho^2 + \rho^{-1}$, where $\rho^2 = x^2 + y^2$. Hence, $U = U(x, y)$ becomes a surface of revolution about the axis $x = 0 = y$. Clearly, $U_x = 0 = U_y$ then holds not only at the five points (16₁)–(16₂) but at every point of the circle $x^2 + y^2 = 1$. Correspondingly, it is seen from (1₂) and (12_I)–(12_{III}^{*}) that all five points (16₁)–(16₂) tend to points of the circle $x^2 + y^2 = 1$, as $\mu \rightarrow 0$.

In what follows, it will again be supposed that $0 < \mu < 1$.

§471. Consider, for any fixed μ , the surface (1₁) in a Cartesian (x, y, U) -space. Then, in the notations introduced at the beginning of §167, the sets $\mathbf{P}_h, \mathbf{Z}_h$ and \mathbf{N}_h represent the sets of those points (x, y) of the (x, y) -plane at which the ordinate U of the surface $U = U(x, y)$ lies above, on or below the ordinate of the plane $U = -h$, respectively, where h is any real number; so that, in particular, \mathbf{Z}_h is the orthogonal projection on the (x, y) -plane of the intersection of the plane $U = -h$ with the surface $U = U(x, y)$, provided that this intersection exists.

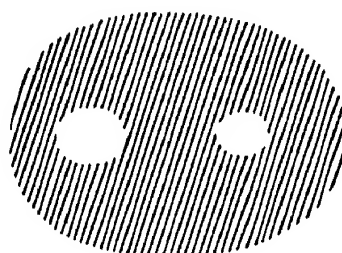
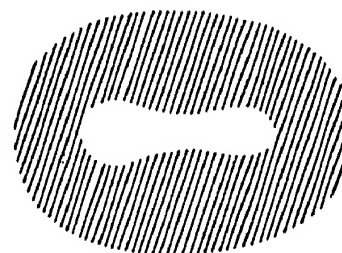
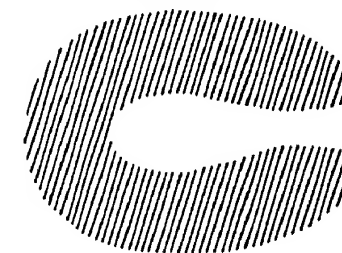
Since the surface is analytic (and, in fact, algebraic), the topological structure of \mathbf{Z}_h , and of the regions $\mathbf{P}_h, \mathbf{N}_h$ into which \mathbf{Z}_h subdivides the (x, y) -plane, cannot change when h varies on an h -interval which is free of h -values of the form $h = -U(a, b)$, where (a, b) is a critical point of the surface, i.e., a point (x, y) at which $\text{grad } U = 0$. According to (i)–(ii), §469, there are exactly five such points (a, b) . Let (a_k, b_k) , where $b_k = 0, a_I < a_{II} < a_{III}$, and $(a_{IV}, b_{IV}), (a_V, b_V)$, where $a_{IV} = a_V, b_{IV} = -b_V$, denote the three collinear and two equilateral critical points, (16₁) and (16₂), respectively. Assuming, without loss of generality, that $\mu \leq 1 - \mu$, and excluding, for sake of convenience, the limiting case $\mu = 1 - \mu$ of two equal masses, one sees from (15₁)–(15₂) and from (iv)–(v) of §469, that $+\infty > U_{II} > U_{III} > U_I > U_{IV} (= U_V = \min U(x, y) > 0)$, where $U_j = U(a_j, b_j)$; $j = I, \dots, V$.

Consequently, the topological structure of $\mathbf{P}_h, \mathbf{N}_h, \mathbf{Z}_h$ does not de-

pend on the value of h as long as $-h$ is within any of the four h -intervals $+\infty > -h > U_{II}$; $U_{II} > -h > U_{III}$; $U_{III} > -h > U_I$; $U_I > -h > U_{IV}$ (the third of which does not exist in the limiting case $\mu = \frac{1}{2}$, where $U_I = U_{III}$). Furthermore, \mathbf{P}_h does not exist if $U_{IV} > -h > -\infty$, the curve \mathbf{Z}_h degenerating into the pair of points (16₂) when $-h$ becomes $U_{IV} = \min U(x, y)$.

§472. Since the locus \mathbf{Z}_h in the (x, y) -plane is defined by the equation $U(x, y) = -h$, one sees from (1₁)–(1₂) that if $-h$ is a large positive number, \mathbf{Z}_h consists of three branches, say \mathbf{B}_h^1 , \mathbf{B}_h^2 , \mathbf{B}_h^3 , the curves \mathbf{B}_h^1 and \mathbf{B}_h^2 being very small, nearly circular curves surrounding the masses (2₁)–(2₂), and \mathbf{B}_h^3 a very large, nearly circular curve about the origin; while the region \mathbf{P}_h in the (x, y) -plane, being defined by the inequality $U(x, y) > -h$, consists of the three disjoint domains which represent the interiors of \mathbf{B}_h^1 and \mathbf{B}_h^2 and the exterior of \mathbf{B}_h^3 , respectively. According to §471, the topological situation is unchanged if $-h$, instead of being very large, merely exceeds the value U_{II} , which belongs to the saddle point of highest ordinate.

On adapting to the present case the considerations of §312, one can readily follow* what happens when $-h$ passes through the successive critical values U_{II} , U_{III} , U_I , U_{IV} ($= U_V$). The situation is schematically illustrated in the four figures, which respectively belong to the four h -intervals mentioned, at the end of §471; the shaded domains representing the regions \mathbf{N}_h , and the boundaries of the shaded regions the curves \mathbf{Z}_h . The third stage disappears in the symmetric case, $\mu = \frac{1}{2}$.

FIG. 14₁FIG. 14₂FIG. 14₃FIG. 14₄

§473. On comparing (6₂), §443 with §167, one sees that $\mathbf{Z}_{-\frac{1}{2}c}$ is the curve of zero velocity belonging to a given value of the energy con-

* The details of the discussion are similar to those given in §496 below.

stant (7₅), §443, and $\mathbf{N}_{-\frac{1}{2}C}$ is the region in the (x, y) -plane which is prohibited for any solution path which belongs to a given value of the Jacobi constant C . If $h = -\frac{1}{2}C$ is less than the positive number $\frac{1}{2}(3 - \mu + \mu^2)$ mentioned at the end of §469, then $\mathbf{N}_{-\frac{1}{2}C}$ contains no point at all (cf. the end of §471); so that the whole (x, y) -plane is then allowed, as far as the energy integral is concerned.

The general results of §167–§170 and §238–§240 are now applicable (and were, as a matter of fact, first found in connection with the restricted problem of three bodies).

§474. It is easy to discuss the equilibrium solutions of the restricted problem of three bodies, that is, the solutions of (6₁), §443 which have the form $x(t) \equiv a = \text{const.}$, $y(t) \equiv b = \text{Const.}$ Clearly, the necessary and sufficient condition for such a pair of constants a, b is that $U_x(x, y) = 0 = U_y(x, y)$ at $(x, y) = (a, b)$. It follows, therefore, from (i)–(ii), §469 that, no matter what the value of μ ($0 < \mu < 1$), there exist exactly five equilibrium solutions, the five pairs $(x, y) = (a, b)$ being represented by (16₁)–(16₂).

Notice that these solutions of equilibrium are the limiting cases, belonging to one body of vanishing mass, of the solutions of relative equilibrium (§380) in the problem of $n = 3$ bodies. In particular, (10) and (7_{III}) represent the three quintic equations obtained from (11), §358 in accordance with the end of §358, if one $m_i = 0$. Similarly, the considerations of §475, §476 will correspond to those of §381, §382, respectively.

§475. If (a, b) denotes any of the five points (16₁)–(16₂), and $\xi = \xi(t)$, $\eta = \eta(t)$ the displacement of the solution $x(t) \equiv a$, $y(t) \equiv b$ of (6₁), §443, then the corresponding Jacobi equations (§86) are seen to be

$$(19) \quad \begin{aligned} \xi'' - 2\eta' &= U_{xx}\xi + U_{xy}\eta, \\ \eta'' + 2\xi' &= U_{xy}\xi + U_{yy}\eta, \end{aligned}$$

where
$$\begin{pmatrix} U_{xx} & U_{xy} \\ U_{xy} & U_{yy} \end{pmatrix} \equiv \begin{pmatrix} U_{xx}(a, b) & U_{xy}(a, b) \\ U_{xy}(a, b) & U_{yy}(a, b) \end{pmatrix} = \text{const.}$$

In order to obtain, in any of the five cases, the four characteristic exponents s by means of the procedure mentioned in §89, one has to determine those numbers s for which (19) admits a solution of the

form $\xi = Ae^{st}$, $\eta = Be^{st}$, where A , B are suitable constants which do not both vanish. Hence, the four s are determined by

$$(20) \quad 0 = \begin{vmatrix} s^2 - U_{xx} & -2s - U_{xy} \\ 2s - U_{xy} & s^2 - U_{yy} \end{vmatrix} \\ \equiv s^4 - (U_{xx} + U_{yy} - 4)s^2 + \begin{vmatrix} U_{xx} & U_{xy} \\ U_{xy} & U_{yy} \end{vmatrix},$$

a quadratic equation in s^2 . Denoting by $(-)$ a certain negative, and by $(?)$ a certain real number (each of which depends on the fixed value of μ), one sees from (17₁)–(17₂) that (20) may be written as

$$(21_1) \quad s^4 + (?)s^2 + (-) = 0; \quad (21_2) \quad s^4 + s^2 + \frac{27}{4}\mu(1 - \mu) = 0,$$

according as the equilibrium solution represented by (a, b) is one of the three collinear points (16₁) or one of the two equilateral points (16₂).

§476. These two cases behave differently, namely as follows:

(I) For any of the three collinear equilibrium solutions (16₁) and for every μ , the four characteristic exponents $s = s(\mu)$ are of the form $s = \pm \alpha$, $s = \pm i\beta$, where α and β are positive functions of μ ; so that the four s are always distinct and never all of the stable type (cf. §89).

(II) For either of the equilateral equilibrium solutions (16₂), three cases are possible, according as the mass contained by one of the two bodies is greater than, less than, or equal to $100(\frac{1}{2} + \frac{1}{18}\sqrt{69})$ percent (about 96%) of the total mass $1 - \mu + \mu = 1$ (so that the two distinct percentages mentioned under (II), §382 coincide in the present problem of a vanishing third mass. In the first case, all four $s = s(\mu)$ are of the stable type and distinct. In the second case, none of the four $s = s(\mu)$ is of the stable type; but all four are, in contrast to the case (I), of the form $s = \pm \alpha \pm i\beta$, where neither of the positive functions α , β of μ vanishes. In the third case, the four s are of the form $s = \pm i\beta_0$, $s = \pm i\beta_0$, where β_0 is both times the same positive number. In this limiting case, the general solution of (19) contains secular terms.

In order to prove (I), it is sufficient to show that one of the roots s^2 of the quadratic equation (21₁) is positive, the other negative. But this is obvious, since the constant term of (21₁) is negative.

In order to prove (II), notice first that both roots s^2 of the quad-

ratic equation (21₂) are negative or both are complex but not purely imaginary, according as the discriminant, $27\mu(1 - \mu) - 1$, is negative or positive. Furthermore, the quadratic condition $27\mu(1 - \mu) - 1 = 0$ for the limiting case is easily verified to be equivalent to the percentual formulation given under (II). Hence, in order to complete the proof of (II), it is sufficient to verify the appearance of secular terms in the limiting case $27\mu(1 - \mu) - 1 = 0$. But such terms then follow from (17₂) by direct integration of (19).

§477. Although the inequality $\min(\mu, 1 - \mu) < 0.03852 \dots$ is, by (II), §476, sufficient (and necessary) for the stable type of the equations of variation belonging to the equilateral solutions (16₂) of equilibrium, §136 bis shows that one cannot be sure of the stability of these solutions, when stability is meant in the sense of §131. Actually, it is to-day an unsolved problem whether these solutions are or are not stable (in the sense of §131). All that can be shown is that, if the answer is affirmative, the stability must be due to the presence of the Coriolis forces. In other words, the solutions (16₂) of the restricted problem of three bodies would certainly not be stable in the sense of §131, if one should omit the terms $-2x', 2y'$ of (6₁), §443.

§477 bis. In order to prove this, consider a point of equilibrium of a reversible dynamical system $x'' = U_x, y'' = U_y$. It may be assumed without loss of generality that this point is the origin $(x, y) = (0, 0)$, and that $U(0, 0) = 0$; so that, since $\text{grad } U(0, 0) = 0$, there exist three constants a, b, c such that

$$(22) \quad \begin{aligned} U(x, y) &= \frac{1}{2}(ax^2 + 2bxy + cy^2) + \dots; \\ U_x &= ax + by + \dots, \quad U_y = bx + cy + \dots, \end{aligned}$$

where the terms \dots are of higher order. Suppose that not only does $U(x, y)$ itself have an isolated minimum at $(x, y) = (0, 0)$, but that the same holds also for its quadratic part, i.e., that $ac - b^2 > 0$ and $a + c > 0$ in (22). It will be shown that this condition (which is, in view of (17₂), satisfied in the problem of §477), is sufficient in order that the equilibrium solution $x(t) \equiv 0, y(t) \equiv 0$ of $x'' = U_x, y'' = U_y$ be not of the stable type in the sense of §131.

First, the assumption imposed on (22) clearly implies the existence of a sufficiently small $\alpha > 0$ such that $xU_x(x, y) + yU_y(x, y) > 0$ at every point (x, y) of the punctured circle $\Gamma(\alpha): 0 < x^2 + y^2 < \alpha^2$.

Furthermore, one can choose α so small that $U(x, y) > U(0, 0)$, i.e. $U(x, y) > 0$, at every point (x, y) of $\Gamma(\alpha)$.

Now suppose, if possible, that the equilibrium solution $x(t) \equiv 0$, $y(t) \equiv 0$ is stable in the sense of §131. Then there exists for every sufficiently small $\epsilon > 0$ a $\delta = \delta_\epsilon$ such that if (x_0, y_0) is any point of $\Gamma(\delta)$, the solution path $x = x(t)$, $y = y(t)$ which is defined by the initial conditions $x(0) = x_0$, $y(0) = y_0$; $x'(0) = 0$, $y'(0) = 0$ exists, and runs within $\Gamma(\epsilon)$, for $-\infty < t < +\infty$. One can, of course, assume that $\delta < \epsilon < \alpha$. But the existence of such a $\delta = \delta_\epsilon$ readily leads to a contradiction.

In fact, for the solution $x = x(t)$, $y = y(t)$ defined by the initial conditions $(x_0, y_0; 0, 0)$, the energy constant $h = \frac{1}{2}(x'^2 + y'^2) - U(x, y)$ reduces to $h = -U(x_0, y_0)$. Hence, the equation of the curve of zero velocity is $U(x, y) = U(x_0, y_0)$, and the solution path can never reach a point (x, y) at which $U(x, y) < U(x_0, y_0)$. Since $U(x, y)$ has at $(x, y) = (0, 0)$ an isolated minimum, and since $(x_0, y_0) \neq (0, 0)$, there follows the existence of a sufficiently small $\eta > 0$ which may be assumed to be less than δ and is such that no point of the solution path is within $\Gamma(\eta)$. Consequently, the solution path is within the ring $\eta^2 \leq x^2 + y^2 \leq \delta^2$ for every t . Since this ring is contained in $\Gamma(\alpha)$, it follows from the definition of α that $xU_x(x, y) + yU_y(x, y)$ has on this ring a positive minimum, say λ . Since the equations $x'' = U_x$, $y'' = U_y$ clearly imply that $(x^2 + y^2)'' = 2(x'^2 + y'^2) + 2(xU_x + yU_y)$, it follows that $(x^2 + y^2)'' \geq 2(x'^2 + y'^2) + 2\lambda$ for every t . Consequently, $(x^2 + y^2)'' \geq 2\lambda = \text{const.} > 0$ for every t . But this implies that $x^2 + y^2 \rightarrow +\infty$, as $t \rightarrow \pm\infty$.

Clearly, an equivalent arrangement of the above proof could have been supplied by a direct verification of the fact that the condition of §133 is not satisfied.

The Non-Planar Restricted Problem

§478. Consider the same model as in §441, but now let the initial position and the initial velocity vector of the third particle, P , be not restricted to the plane of the circular motion of the two finite bodies P_1, P_2 ; so that P is not required to move in this plane and has, therefore, three, instead of two, degrees of freedom. It is clear that (5₁)–(5₂), §443 must then be replaced by

$$(1_1) \quad L = \frac{1}{2}(x'^2 + y'^2 + z'^2) + (xy' - yx') + U(x, y, z),$$

$$(1_2) \quad U = \frac{1}{2}(x^2 + y^2) + \frac{1 - \mu}{|(x + \mu)^2 + y^2 + z^2|^{\frac{1}{2}}} + \frac{\mu}{|(x - 1 + \mu)^2 + y^2 + z^2|^{\frac{1}{2}}},$$

where (x, y, z) denote the barycentric synodical coordinates of P , and the rotating (x, y) -plane coincides with the non-rotating plane which contains the circular paths of P_1 and P_2 . The three Lagrangian equations determined by (1₁)-(1₂) are seen to be

$$(2_1) \quad x'' - 2y' = U_x; \quad (2_2) \quad y'' + 2x' = U_y; \quad (2_3) \quad z'' = U_z.$$

The energy integral is

$$(3) \quad \frac{1}{2}(x'^2 + y'^2 + z'^2) - U(x, y, z) = \text{const.}$$

§479. One is led to an elementary type of motion by requiring that $x(t) \equiv 0$ and $y(t) \equiv 0$, i.e., that the motion of P take place along the z -axis.

For a motion of this kind, (2₁), (2₂) require that $0 = U_x(0, 0, z)$, $0 = U_y(0, 0, z)$. This is, in view of (1₂), equivalent to $\mu = 1 - \mu$, since $0 < \mu < 1$; so that P_1 and P_2 have the same mass $\mu = \frac{1}{2}$ and, therefore, the coordinates $(x, y, z) = (\pm \frac{1}{2}, 0, 0)$ for every t . Since P is supposed to move along the z -axis, it follows that the triangle formed by the three bodies must be isosceles for every t .

In order to find the ordinate $z = z(t)$ of P , one has to satisfy (2₃) by $x(t) \equiv 0$, $y(t) \equiv 0$ and $\mu = \frac{1}{2}$. Thus, $z'' = U_z = -z/(z^2 + \frac{1}{4})^{\frac{3}{2}}$, by (1₂). This is a dynamical system with a single degree of freedom, admitting the energy integral $\frac{1}{2}z'^2 - U(z) = \text{const.}$; whence $z = z(t)$ follows by the inversion of a quadrature (leading to an elliptic function).

§480. Let $x = x(t)$, $y = y(t)$, $z = z(t)$ be any solution path which is neither of the type $x(t) \equiv 0 \equiv y(t)$, considered in §479, nor of the type $z(t) \equiv 0$, considered before §478. Suppose that

$$(4) \quad x(t)y'(t) - y(t)x'(t) \neq 0,$$

and let $P = P(t)$ denote, at a given t , the osculating plane of the curve represented by the solution path in the (x, y, z) -space. The Eulerian angles of the plane $P = P(t)$ with reference to the (x, y) -plane will be denoted by $\iota = \iota(t)$ and $\vartheta = \vartheta(t)$, respectively; ι denoting the inclination of P and ϑ its node, that is to say the angle be-

tween the x -axis and the line along which the plane P cuts the (x, y) -plane (if $\sin \iota \neq 0$).

On choosing the orientation of these angles in a suitable way, and denoting by $R = R(t)$ the vector product of the vectors $(x(t), y(t), z(t))$ and $(x'(t), y'(t), z'(t))$, one has

$$\begin{aligned} (5) \quad yz' - zy' &= |R| \sin \iota \sin \vartheta, \\ zx' - xz' &= -|R| \sin \iota \cos \vartheta, \\ xy' - yx' &= |R| \cos \iota. \end{aligned}$$

For, on the one hand, $-\sin \iota \sin \vartheta, \sin \iota \cos \vartheta, -\cos \iota$ are seen to be the direction cosines which determine the position of the normal of P with reference to the respective axes x, y, z ; and, on the other hand, R is, by the definition of P as osculating plane, perpendicular to P .

If $\cos \iota \neq 0$, then (5) implies, by (4), that

$$(6) \quad z = (-x \sin \vartheta + y \cos \vartheta) \tan \iota, \quad z' = (-x' \sin \vartheta + y' \cos \vartheta) \tan \iota.$$

§481. The solutions found in §479 have no astronomical interest. Relevant for the applications is the other extreme case, that in which the particle moves in a region close to the (x, y) -plane; so that $z = z(t)$, without vanishing identically, is very small in absolute value.

In order to deal with this situation, suppose that there is given a planar solution

$$(7) \quad x = x(t), \quad y = y(t), \quad (z = z(t) \equiv 0),$$

of (2₁), (2₂), (2₃), i.e., of (2₁)–(2₂). Non-planar solutions which are very close to this planar solution may be obtained approximately by replacing (2₃) by its Jacobi equation with reference to (7). In fact, if $\zeta = \zeta(t)$ denotes the displacement of $z(t) \equiv 0$ in the sense of §86, then the Jacobi equation is obtained by neglecting on the right of (2₃) all terms which are not of the first order in z , and then writing $x(t), y(t), \zeta$ for x, y, z , respectively. Thus, the approximate or Jacobi differential equation which determines the ordinate is

$$(8) \quad \zeta'' = -f(t)\zeta, \quad \text{where } -f(t) = U_{zz}(x(t), y(t), 0); \text{ cf. (1}_2\text{) and (7).}$$

And if $\zeta = \zeta(t)$ is any solution of the linear differential equation (8), then, unless $\zeta(t) \equiv 0$, an approximate non-planar solution of (2₁)–(2₃) is represented by $x = x(t), y = y(t), z = \zeta(t)$, where the precise

meaning of the adjective "approximate" is sufficiently clear from §84–§86 (cf. also §136).

§482. Clearly, the inclination $\iota = \iota(t)$ considered in §480 must be very small under the assumption made in §481; so that, on replacing z by ζ , one can replace (6) by

$$(9) \quad \zeta = (-x \sin \vartheta + y \cos \vartheta)\iota, \quad \zeta' = (-x' \sin \vartheta + y' \cos \vartheta)\iota,$$

where x, y , hence also x', y' , are functions of t which are given by (7).

It follows that, barring the trivial solution $\zeta(t) \equiv 0$ of (8), one can replace the differential equation (8) of the second order for ζ by a system of two differential equations of the first order for ϑ, ι . In fact, it is clear from (9) and (4) that the Jacobian of (ζ, ζ') with respect to (ϑ, ι) vanishes if and only if so does ι . But if $\iota = \iota(t)$ vanishes at some $t = t_0$, then so do ζ and ζ' , by (9). And the solution of (8) which belongs to the initial condition $\zeta(t_0) = 0, \zeta'(t_0) = 0$ is $\zeta(t) \equiv 0$.

§483. In order to obtain the explicit representation of (8) in terms of the Eulerian angles ϑ, ι , it is convenient to replace ϑ, ι by their combinations

$$(10) \quad u = (xy' - yx')^{\frac{1}{2}} \iota \cos \vartheta, \quad v = (xy' - yx')^{\frac{1}{2}} \iota \sin \vartheta,$$

if the given non-vanishing continuous function (4) of t is positive, and to modify (10) in an obvious manner, if (4) is negative. According to (10), one can write (9) in the form

$$(11) \quad p = (xy' - yx')^{-\frac{1}{2}}(yu - xv), \quad q = (xy' - yx')^{-\frac{1}{2}}(y'u - x'v),$$

if one puts $\zeta = p, \zeta' = q$. But (11) is a linear transformation of (u, v) into (p, q) , with coefficients which are, by (7), given functions of t , and have the determinant 1 for every t . It follows, therefore, from §40 that (11) is a canonical transformation of multiplier 1. On the other hand, (8) may be written as a linear canonical system with a single degree of freedom, the Hamiltonian function being the quadratic form $H(p, q; t) = -\frac{1}{2}q^2 - \frac{1}{2}f(t)p^2$. Thus, on subjecting this system to the linear canonical transformation (11), one obtains for u, v a linear canonical system having as Hamiltonian function the quadratic form $K(u, v; t) = H$ plus a remainder function. Finally, the explicit form of this remainder function follows from (11) and (4) by the rule (17₁)–(17₂), §38.

§483 bis. In the theory of the Moon, that case of the Jacobian equation (8), or of the equivalent canonical system, is of particular interest in which the underlying planar solution (7) is periodic (cf. §517 below). This is the case which will be analyzed in what follows.

§484. The treatment will be based on a theorem concerning complex-valued functions $u + iv \equiv w = w(t)$ of a real variable t which are almost periodic in the sense of H. Bohr. Suppose that such a $w(t)$ satisfies the condition $|w(t)| > \text{const.}$ for some $\text{const.} > 0$ and for $-\infty < t < +\infty$; so that $w(t)/|w(t)|$ is an almost periodic function of absolute value 1 for all t , and has frequencies which are all contained in the integral modul of the frequencies of $w(t)$. Put $w(t)/|w(t)| = \exp i\vartheta(t)$; so that $\vartheta(t)$ is a real function which may be chosen to be continuous and is then uniquely determined by the normalization $0 \leq \vartheta(0) < 2\pi$, say. Thus, $\vartheta(t) = \arg w(t)$, where $w = u + iv$; so that $(u^2 + v^2)^{\frac{1}{2}}$ and ϑ are polar coordinates in the (u, v) -plane. Then the theorem to be used states that there exist a unique real constant ω and a unique real almost periodic function $\psi(t)$ for which $\vartheta(t) = \omega t + \psi(t)$; and that ω and the frequencies of $\psi(t)$ are contained in the integral modul of the frequencies of the almost-periodic function $\exp i\vartheta(t) = w(t)/|w(t)|$ (and so in that of the function $w(t)$). The proof of this known general theorem will not be reproduced here.

The coefficient ω of the "secular" part ωt of $\vartheta(t)$ is called the mean motion* of $\vartheta(t)$. It is understood that ω may accidentally vanish as may ψ . In fact, it is clear that if $\psi(t)$ is any real almost periodic function and ω any real constant, then $\exp i\vartheta(t)$, where $\vartheta(t) = \omega t + \psi(t)$, is almost periodic.

§485. Consider any linear system

$$(15) \quad u' = a_{11}(t)u + a_{12}(t)v, \quad v' = a_{21}(t)u + a_{22}(t)v$$

in which the given coefficient functions $a(t)$ are real, continuous, and

* The origin of this name is that if $\vartheta = \vartheta(t)$ is absolutely continuous and $\vartheta(t):t$ tends, as $t \rightarrow \infty$, to a limit ω , then, since

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \vartheta'(t) dt = \lim_{T \rightarrow \infty} \frac{\vartheta(T) - \vartheta(0)}{T} = \lim_{T \rightarrow \infty} \frac{\vartheta(T)}{T} = \omega,$$

ω represents the average velocity of the angle $\vartheta(t)$; and that, in the terminology of the 17th and 18th centuries, *motio* meant velocity. Thus, "mean motion" = "average velocity."

such as to have a common period, say τ . Suppose that the two characteristic exponents of (15) are of the stable type, i.e., of the form $\pm i\sigma$, where σ is a real constant, determined by the $a(t)$; and that the monodromy group has no multiple elementary divisors. Thus, by §144, the general solution of (15) is of the form

$$(16) \quad u = C_1 A_{11}(t) e^{i\sigma t} + C_2 A_{12}(t) e^{-i\sigma t}, \quad v = C_1 A_{21}(t) e^{i\sigma t} + C_2 A_{22}(t) e^{-i\sigma t},$$

where the four $A(t)$ have the period τ and are independent of the integration constants C_1, C_2 . Since only real solutions of (15) are to be considered, one readily finds (by taking the real parts of the products $C A(t) e^{\pm i\sigma t}$) that (16) is equivalent to a matrix relation

$$(17) \quad \begin{pmatrix} u(t) \\ v(t) \end{pmatrix} = \begin{pmatrix} \alpha_{11}(t) & \alpha_{12}(t) \\ \alpha_{21}(t) & \alpha_{22}(t) \end{pmatrix} \begin{pmatrix} \cos \sigma t & -\sin \sigma t \\ \sin \sigma t & \cos \sigma t \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix},$$

where the $\alpha(t)$ are real, of period τ , and independent of the real integration constants c_1, c_2 . Notice that the matrix product (17) is not, in general, periodic, since the two positive numbers $2\pi/\tau, \sigma$ need not be commensurable. Nevertheless, the solution vector (17) cannot come arbitrarily close to 0 as $t \rightarrow \pm \infty$, if one excludes the trivial solution $u(t) \equiv 0, v(t) \equiv 0$ (which belongs to $c_1 = 0, c_2 = 0$).

In order to prove this, notice first that the second matrix factor on the right of (17) represents a mere rotation of the vector (c_1, c_2) and may, therefore, be disregarded, as far as only the length $(u^2 + v^2)^{\frac{1}{2}}$ of the vector (u, v) is concerned. On the other hand, the first matrix factor on the right of (17) is a continuous periodic function of t . Hence, $u^2(t) + v^2(t)$ has, for $-\infty < t < +\infty$, a lower bound which is the product of $c_1^2 + c_2^2$ and of a number β which is positive or zero according as the continuous periodic function $\det \alpha_{nm}(t)$ does not or does vanish for a suitable fixed t_0 . But the second case is impossible. In fact, $\det \alpha_{nm}(t)$ is identical with the determinant of the fundamental matrix which is the product of the two matrices on the right of (17); so that, by §138, one must have $\det \alpha_{nm}(t) \neq 0$ for every t .

Consequently, there exists a constant $\beta > 0$ such that

$$u^2(t) + v^2(t) \geq (c_1^2 + c_2^2)\beta.$$

§486. It follows that if $u = u(t), v = v(t)$ is any (real) solution of (15) distinct from $u(t) \equiv 0, v(t) \equiv 0$, the condition $|w(t)| > \text{const.} > 0$ of §484 is satisfied by $w(t) = u(t) + iv(t)$. Thus, the polar angle $\vartheta = \vartheta(t)$ in the Cartesian (u, v) -plane admits a decomposition

$\vartheta(t) = \omega t + \psi(t)$ into secular and almost periodic components ωt , $\psi(t)$. Furthermore, the mean motion ω and the frequencies of $\psi(t)$ are homogeneous linear combinations, with integral coefficients, of the two numbers $2\pi/\tau$, σ (which may, but need not, be commensurable). In fact, (17) shows that the same holds for the frequencies of $w(t) = u(t) + iv(t)$, since the $\alpha(t)$ have the period τ .

§487. Returning to the problem of §482–§483 bis, one has to identify (15) with the canonical system found at the end of §483. Thus, the angle $\vartheta = \vartheta(t)$ defined in §484 by $u = (u^2 + v^2)^{\frac{1}{2}} \cos \vartheta$, $v = (u^2 + v^2)^{\frac{1}{2}} \sin \vartheta$ becomes identical with the angle $\vartheta = \vartheta(t)$ defined by (10), §483 (where $(u^2 + v^2)^{\frac{1}{2}} = \pm (xy' - yx')^{\frac{1}{2}} i$). Accordingly, the result of §486 concerns the node $\vartheta = \vartheta(t)$ considered in §480–§483 bis.

§488. It may be mentioned that, from the formal point of view, the problem of integrating (15) is simplified by the introduction of the polar coordinates $\vartheta = \arctan v/u$, $r = (u^2 + v^2)^{\frac{1}{2}}$.

In fact, since $uv' - vu' = r^2\vartheta'$ and $uu' + vv' = rr'$, one can write (15) in the form

$$(18_1) \quad \vartheta' = a_{21}(t) \cos^2 \vartheta + \{a_{22}(t) - a_{11}(t)\} \cos \vartheta \sin \vartheta - a_{12}(t) \sin^2 \vartheta,$$

$$(18_2) \quad (\log r)' = a_{11}(t) \cos^2 \vartheta + \{a_{12}(t) + a_{21}(t)\} \cos \vartheta \sin \vartheta + a_{22}(t) \sin^2 \vartheta.$$

Notice that the function on the right of each of these differential equations is a continuous function of the position on a (ϑ, t) -torus, since it is periodic in both ϑ and t . If a solution $\vartheta = \vartheta(t)$ of the differential equation (18₁) is known, $r = r(t)$ follows from (18₂) by a quadrature. Incidentally, (18₁) may be written as a Riccati differential equation for $e^{2i\vartheta}$.

It is readily seen that if (15) is a canonical system, say

$$u' = -K_v, \quad v' = K_u,$$

where $K = K(u, v; t)$ is a quadratic form in u, v , then (18₁) reduces to $\vartheta' = 2K(\cos \vartheta, \sin \vartheta; t)$.

Lunar Systems

§489. According to §443, the restricted problem of three bodies may be defined by

$$(1_1) \quad x'' - 2y' = U_x, \quad y'' + 2x' = U_y;$$

$$(1_2) \quad U = \frac{1}{2}(x^2 + y^2) + \frac{1 - \mu}{|(x + \mu)^2 + y^2|^{\frac{1}{2}}} + \frac{\mu}{|(x + \mu - 1)^2 + y^2|^{\frac{1}{2}}},$$

where the origin of the rotating coordinate system (x, y) is the centre of mass of the bodies $\mu, 1 - \mu$, which rest at the respective points $(x, y) = (1 - \mu, 0), (x, y) = (-\mu, 0)$. Hence, if the origin of the coordinate system is transferred to the mass μ , the coordinates x, y of the third body must be replaced by $\xi = x - (1 - \mu), \eta = y$. Substituting the inverse of this transformation into (1₁)–(1₂) and then writing x, y for ξ, η , one sees from $1 - \mu = \text{const.}$ that (1₁) is again valid if (1₂) is replaced by

$$(2) \quad U = \frac{1}{2}(1 - \mu)^2 + (1 - \mu)x + \frac{1}{2}(x^2 + y^2) \\ + (1 - \mu)(1 + 2x + x^2 + y^2)^{-\frac{1}{2}} + \mu(x^2 + y^2)^{-\frac{1}{2}}.$$

The masses $\mu, 1 - \mu$ now rest at the respective points

$$(x, y) = (0, 0), \quad (x, y) = (-1, 0).$$

The following considerations are relevant if the orders of magnitude involved are such as those in the case in which $\mu: (0, 0)$ signifies the Earth, $1 - \mu: (-1, 0)$ the Sun, and the third body the Moon: (x, y) .

§490. Suppose that the third body moves in a region whose points are rather close to the permanent position $(0, 0)$ of the body μ ; and that, correspondingly, one wishes to neglect in (1₁) all terms which are at least of the second order in $(x^2 + y^2)^{\frac{1}{2}}$, that is, in $|x|$ and $|y|$ together. This means the rejection of those terms of (2) which are at least of the third order in $|x|$ and $|y|$ together. Then, by the binomial expansion,

$$(1 + 2x + x^2 + y^2)^{-\frac{1}{2}} = 1 - \frac{1}{2}(2x + x^2 + y^2) + \frac{3}{8}(2x + \dots)^2 - \dots$$

Substituting this into (2), one sees that, to the required degree of approximation,

$$(3) \quad U = \text{const.} + \frac{1}{2}\mu(x^2 + y^2) + \frac{3}{2}(1 - \mu)x^2 + \mu(x^2 + y^2)^{-\frac{1}{2}}; \\ \text{const.} = \frac{1}{2}(1 - \mu)^2 + 1 - \mu.$$

§490 bis. In §489–§490, the units were those chosen in §441. And this choice of the units depended on Kepler's third law (cf. §276). Since this law loses its validity by the passage from (1₂) or (2) to the approximation (3), it will now be necessary to obtain direct information on the orders of magnitude involved.

To this end, change the units of distance, mass and time in the respective proportions $1:\alpha, 1:\beta$ and $1:1$, where α and β are arbitrary

positive constants. In other words, substitute (3) into (1₁) and then write αx , αy ; $\beta\mu$, $\beta(1 - \mu)$ instead of x , y ; μ , $1 - \mu$, respectively. On dividing the resulting equations by α , one clearly obtains

$$\begin{aligned}x'' - 2y' &= \beta\mu x + 3\beta(1 - \mu)x - \alpha^{-3}\beta x(x^2 + y^2)^{-\frac{1}{2}}, \\y'' + 2x' &= \beta\mu y - \alpha^{-3}\beta y(x^2 + y^2)^{-\frac{1}{2}}.\end{aligned}$$

§491. These equations are, of course, equivalent to (1₁) in the case (3); so that the assumption in §490 bis is the same as in §490, namely, that the third body moves in a region rather close to the first body, the latter resting at $(x, y) = (0, 0)$ and having, in terms of the present units, the mass $\beta\mu$. Now suppose that this mass $\beta\mu$ is very small when compared with the mass $\beta(1 - \mu)$ of the second body. Then, $|x|$ and $|y|$ being small by the assumption of §490, a close approximation to the equations given at the end of §490 bis is represented by

$$\begin{aligned}x'' - 2y' &= 3\beta(1 - \mu)x - \alpha^{-3}\beta x(x^2 + y^2)^{-\frac{1}{2}}, \\y'' + 2x' &= -\alpha^{-3}\beta y(x^2 + y^2)^{-\frac{1}{2}}.\end{aligned}$$

Clearly, the latter equations may be written in the form (1₁), if one puts $U = \frac{3}{2}\beta(1 - \mu)x^2 + \alpha^{-3}\beta(x^2 + y^2)^{-\frac{1}{2}}$; so that

$$(4) \quad U = \frac{3}{2}x^2 + (x^2 + y^2)^{-\frac{1}{2}},$$

if the units of distance and mass, which in §490 bis were made arbitrary by the introduction of the factors α and β , are now determined by the conditions $\alpha = (1 - \mu)^{-\frac{1}{2}}$ and $\beta = (1 - \mu)^{-1}$.

§492. On using the interpretation given at the end of §489, one can say, that, in view of Kepler's third law, the assumption which in §491 was added to that of §490 is equivalent to the assumption that the distance Earth-Sun is very large; while the assumption of §490 is that the distance Moon-Earth is relatively small. Hence, on using the terminology of lunar theory, one can say that the approximation (4) to (3) does, while the approximation (3) to (1₂) does not, neglect the parallax; and that the transition from (1₂) to (3) neglects the second and the higher powers of this parallax. It is understood that the parallax may, roughly, be defined as the ratio of the distances Moon-Earth and Earth-Sun.

§493. Actually, (1₁) with (4) represents the foundation of the modern theory of the Moon. This problem of two degrees of freedom is called Hill's limiting case of the restricted problem of three bodies.

From the analytical point of view, there is hardly a difference between the two problems (1₁) defined by (4) and by (3). On the other hand, the only formal difference between (3) and (2) is that, while (2) has the two singularities $(x, y) = (0, 0)$ and $(x, y) = (-1, 0)$, only the first of these singularities appears in (3). And most of the principal mathematical problems of a general nature which arise for (1₁) in the cases (1₂) or (2) arise also in the case (4). In this case,

$$(5_1) \quad x'' - 2y' = U_x, \quad y'' + 2x' = U_y;$$

$$(5_2) \quad \frac{1}{2}(x'^2 + y'^2) - U(x, y) = -\frac{1}{2}C,$$

(5₂) being the energy integral of (5₁) in terms of the constant C , which is called, as in the case of §443, the Jacobi constant.

The fact that the function (4) which occurs in (5₁)–(5₂) has only one singular point (that at the position $(x, y) = (0, 0)$ of the Earth) enables one to eliminate between (5₁) and (5₂) this singularity which is represented by the term $(x^2 + y^2)^{-\frac{1}{2}}$ and its partial derivatives. In fact, one readily finds from (4) that (5₁) may be written in virtue of (5₂) in the form

$$(6) \quad \begin{aligned} xy'' - yx'' + 2xx' + 2yy' + 3xy &= 0, \\ xx'' + yy'' + 2yx' - 2xy' + \frac{1}{2}x'^2 + \frac{1}{2}y'^2 - \frac{9}{2}x^2 + \frac{1}{2}C &= 0, \end{aligned}$$

which is free of singularities; and (5₂) is an invariant relation (§80) not only of (5₁) but also of (6).

§494. It is clear from the deduction of (4) that, in this limiting case of (2), the large mass $1 - \mu$ may be thought of as being situated at $x = -\infty$ on the axis $y = 0$ of syzygies.

It is indicated by this remark that the third of the collinear and both of the equilateral points (16₁)–(16₂), §469 disappear. Actually, it is readily found from (4) that $U_x = 0 = U_y$ only at the pair of points $(x, y) = (\pm 3^{-\frac{1}{2}}, 0)$, which obviously correspond to the first two of the three points (16₁), §469. The function $U(x, y)$ has a saddle point at each of these points, since the Hessian matrix of (4) at $(x, y) = (\pm 3^{-\frac{1}{2}}, 0)$ is readily found to be of the form (17₁), §469, with $+$ = 9, $-$ = -3.

Substituting these values into (19)–(20), §475, one sees that (21₁), §475 holds with $(?) = -2$, $(-) = -27$; so that the four s belonging to either of the existing equilibrium solutions $x(t) \equiv \pm 3^{-\frac{1}{2}}$, $y(t) \equiv 0$ are $s = \pm \sqrt{1 + 2\sqrt{7}}$, $s = \pm i\sqrt{-1 + 2\sqrt{7}}$. This agrees with (I), §476.

§495. The function (4), which is everywhere positive, tends to 0 as $x^2 + y^2 \rightarrow +\infty$ along the y -axis. On the other hand, (4) tends to $+\infty$ as $x^2 + y^2 \rightarrow +\infty$ along any half-line not on the y -axis, and this holds uniformly for every closed set of such half-lines in the (x, y) -plane. In addition, (4) becomes $+\infty$ at $(x, y) = (0, 0)$. Furthermore, the surface $U = U(x, y)$ in an (x, y, U) -space is, by (4), symmetric with respect to both planes $x = 0, y = 0$. According to §494, the tangent plane to this surface is parallel to the (x, y) -plane only at the two points $(x, y) = (\pm 3^{-\frac{1}{3}}, 0)$, and the surface has at these points saddle points, the Hessian matrix being indefinite. The ordinate U at these saddle points is $\frac{2}{3}\sqrt[3]{3}$, by (4).

§496. As in §471, let $\mathbf{P}_h, \mathbf{Z}_h$ and \mathbf{N}_h denote the sets of those points of the (x, y) -plane at which the ordinate U of the surface $U = U(x, y)$ lies above, on or below the ordinate of the plane $U = -h$, respectively, where h is a fixed real number.

If $0 \leq h < +\infty$, then \mathbf{P}_h and \mathbf{Z}_h contain no point (i.e., \mathbf{N}_h is the whole plane), since (4) is everywhere positive.

If $-\infty < h < 0$, the topological structure of the regions $\mathbf{P}_h, \mathbf{N}_h$ and of their border \mathbf{Z}_h depends only on whether $-h$ is less than, greater than, or equal to the critical value $\frac{2}{3}\sqrt[3]{3}$. In fact, $\frac{2}{3}\sqrt[3]{3}$ is, by §495, the only value which is ordinate of critical points of the surface, i.e., of points at which the tangent plane is parallel to the (x, y) -plane ($\text{grad } U = 0$).

In all three cases $-h \leq \frac{2}{3}\sqrt[3]{3}$ possible for $-\infty < h < 0$, one sees from §495 that the curve $\mathbf{Z}_h: U(x, y) = -h$ is symmetric with respect to both coordinate axes x, y , and must possess asymptotes parallel to the y -axis. In view of (4), these asymptotes are the two lines $x = \pm (-\frac{2}{3}h)^{\frac{1}{3}}$. It is also seen from (4) that, in all three cases possible for $-\infty < h < 0$, the curve \mathbf{Z}_h intersects the y -axis at the two points $(x, y) = (0, \pm h^{-1})$; while the pair of points $(\pm |x_0|, 0)$ of \mathbf{Z}_h on the x -axis is determined by the cubic equation $|x_0|^3 + \frac{2}{3}h|x_0| + \frac{2}{3} = 0$ for $|x_0| > 0$. Since the discriminant, $-4(\frac{2}{3}h)^3 - 27(\frac{2}{3})^2$, is of the same sign as $-h - \frac{2}{3}\sqrt[3]{3}$, it follows that the number of the pairs $(\pm |x_0|, 0)$ of points of \mathbf{Z}_h on the x -axis is 0, 1, 2 according as the given positive parameter $-h \leq \frac{2}{3}\sqrt[3]{3}$.

§497. These three cases are schematically illustrated in the figures* in which C denotes $-2h$ and the (real) branches of the alge-

* Only Fig. 15_I and Fig. 15_{II} correspond to Fig. 14_I–Fig. 14₄ (§472), since Fig. 15_{II} corresponds to the three limiting cases which form the transitions

braic curve $Z_{-\frac{1}{2}C}$: $U(x, y) = \frac{1}{2}C$ are represented by the boundary between the shaded and unshaded regions, the latter being $P_{-\frac{1}{2}C}$ and $N_{-\frac{1}{2}C}$, respectively.

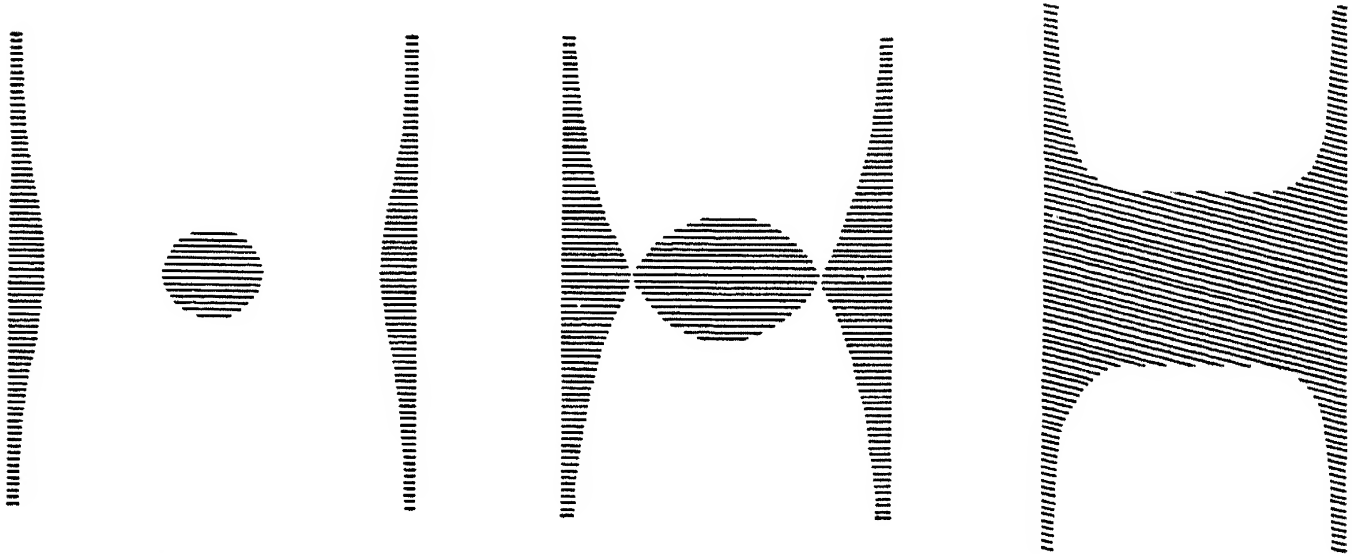


FIG. 15I

FIG. 15II

FIG. 15III

In view of (5₂), the branches of $Z_{-\frac{1}{2}C}$, where $0 < C < +\infty$, represent the curve of zero velocity belonging to a fixed energy $h = -\frac{1}{2}C$, while the unshaded regions, $N_{-\frac{1}{2}C}$, are those precluded by the energy integral (cf. §167).

It is clear from §492 that, in view of the assumptions which underlie the replacement of (2) by (4), only that case is astronomically significant in which the path $x = x(t)$, $y = y(t)$ of the Moon: (x, y) is assured of running in the neighborhood of the Earth: $(0, 0)$. According to Fig. 15I–Fig. 15III, this will be the case only if C has a large value and, in addition, the initial position (x_0, y_0) of the Moon is chosen within that one of the three shaded regions of Fig. 15I which is the bounded component of $P_{-\frac{1}{2}C}$ (i.e., in the shaded region surrounding the origin). This region is, in view of (4), approximately represented by the interior of the circle $(x^2 + y^2)^{-\frac{1}{2}} = \frac{1}{2}C$ of radius $2C^{-1}(\rightarrow 0)$ about the Earth.

§498. In order to regularize (5₁)–(5₂) for any fixed C , where $-\infty < C < +\infty$, one can replace (6) by the equations which re-

between the four stages represented in Fig. 14I–Fig. 14₄ (the limiting cases are not illustrated by the figures of §472). Needless to say, the shaded regions represent the admissible domains in Fig. 15I–Fig. 15III and the prohibited domains in Fig. 14I–Fig. 14₄.

sult by applying to the present force function (4) the transformation of §446 in case of the parabolic mapping

$$(7) \quad x + iy \equiv z = \zeta^2 \equiv (\xi + i\eta)^2, \text{ i.e., } x = \xi^2 - \eta^2, y = 2\xi\eta; \text{ cf. §54.}$$

In fact, (9₁)–(9₂), §446 then become

$$(8_1) \quad \dot{\xi}^2 + \dot{\eta}^2 = 2\bar{U}(\xi, \eta; -\tfrac{1}{2}C);$$

$$(8_2) \quad \bar{U} = 4 - 2(\xi^2 + \eta^2)C + 6(\xi^2 - \eta^2)^2(\xi^2 + \eta^2);$$

while (8₁)–(8₂), §446 reduce, corresponding to (10₁)–(10₂), §447, to

$$(9_1) \quad \ddot{\xi} - 8(\xi^2 + \eta^2)\dot{\eta} = \bar{U}_{\xi}, \quad \ddot{\eta} + 8(\xi^2 + \eta^2)\dot{\xi} = \bar{U}_{\eta};$$

$$(9_2) \quad \dot{t} = 4(\xi^2 + \eta^2),$$

the dots denoting differentiations with respect to the time variable $\bar{t} = \bar{t}(t)$ which follows from (9₂) by the inversion of a quadrature.

Comparison of (8₂) with (12), §447 shows that the formulae of §448 remain valid if one puts $\mu = 0$. Thus, if the collision of the Moon: $(\xi(t), \eta(t))$ with the Earth: $(\xi, \eta) = (0, 0)$ takes place when $t = 0$, and if the origin of the \bar{t} -axis is chosen so as to belong to $t = 0$, then, by (13)–(16), §448,

$$(10_1) \quad \xi = (8^{\frac{1}{2}} \cos \gamma) \cdot \bar{t} + \dots, \quad \eta = (8^{\frac{1}{2}} \sin \gamma) \cdot \bar{t} + \dots;$$

$$(10_2) \quad t = \frac{3}{8} \bar{t}^3 + \dots,$$

where γ is an integration constant and $\bar{t} \geq 0$ is sufficiently small. The consequences drawn in §448–§450 from this uniformization of the singularity at the date $t = 0$ of a collision remain valid without change. And the considerations of §455–§461 may be repeated with obvious modifications (and simplifications).

§499. In view of §180 and §231 bis, the connection between (5₁) and (9₁) is to the effect that, barring the pair $x(t) \equiv \pm 3^{-\frac{1}{2}}, y(t) \equiv 0$ of equilibrium solutions (§494), those solutions $x = x(t), y = y(t)$ of (5₁) which belong to a fixed value of the energy constant (5₂) are, in virtue of (7) and (9₂), identical with those solutions $\xi = \xi(\bar{t}), \eta = \eta(\bar{t})$ of (9₁) which satisfy the invariant relation (8₁) of (9₁). In other words, instead of considering the four-dimensional $(\xi, \dot{\eta}, \xi, \eta)$ -space (or, what amounts to the same thing, the phase space in the sense of §16), one considers the three-dimensional isoenergetic hypersurface $\mathbf{F} = \mathbf{F}_C$ which belongs to any fixed value of C and has, in terms of the coordinates of the underlying four-dimensional space, the equation (8₁); so that, from (8₂),

$$(11) \quad \mathbf{F}_C: \dot{\xi}^2 + \dot{\eta}^2 + 4(\xi^2 + \eta^2) \{C - 3(\xi^2 - \eta^2)^2\} = 8.$$

Inasmuch as the parabolic mapping (7) is such that (ξ, η) and $(-\xi, -\eta)$ belong to the same (x, y) , it is understood that the definition of the "points" of the hypersurface \mathbf{F}_C is meant with the proviso that if $(\dot{\xi}, \dot{\eta}, \xi, \eta)$ is a point of \mathbf{F}_C , then $(-\dot{\xi}, -\dot{\eta}, -\xi, -\eta)$ represents the same point of \mathbf{F}_C .

Since the initial values of the velocities may be chosen arbitrarily, it is clear from (7) that the isoenergetic three-dimensional manifold \mathbf{F}_C in the four-dimensional $(\dot{\xi}, \dot{\eta}, \xi, \eta)$ -space consists of as many disconnected parts (components) as does the two-dimensional (x, y) -region which in §496 was denoted by \mathbf{P}_h , where $h = -\frac{1}{2}C$.

§500. Assume the case $3^{\frac{1}{3}} < C < +\infty$ of Fig. 15_I, and denote by \mathbf{F}_C^* that of three components of \mathbf{F}_C which is astronomically significant in the sense explained in §497. Then the manifold of the possible isoenergetic states $(\dot{\xi}, \dot{\eta}, \xi, \eta)$ of motion which are represented by the points of \mathbf{F}_C^* is topologically equivalent to the (real) three-dimensional projective space.

In order to prove this, notice first the topological structure of \mathbf{F}_C^* is independent of C for every $C > 3^{\frac{1}{3}}$. This readily follows, either from the corresponding remark of §472 in view of the critical value $\frac{2}{3}\sqrt[3]{3}$ (§495–§496) of $-h = \frac{1}{2}C$, or, more directly, from an inspection of the rank of the matrix of the partial derivatives of (11). Thus, instead of assuming that C has a fixed value $> 3^{\frac{1}{3}}$, one can assume that the fixed value of C is a large positive number. But in the latter case the last remark of §497 is applicable and implies, in view of (7), that the set of those points of the (ξ, η) -plane for which the point $(\dot{\xi}, \dot{\eta}, \xi, \eta)$ of the four-dimensional space is a point of the three-dimensional manifold \mathbf{F}_C^* for suitable $(\dot{\xi}, \dot{\eta})$, consists of a simply connected domain which is approximately represented by the small circle $\xi^2 + \eta^2 \leq 2C^{-1}$ ($\rightarrow 0$) about the origin of the (ξ, η) -plane. Thus, C being a large positive constant, $|\xi|$ and $|\eta|$ are small uniformly for all points of \mathbf{F}_C^* ; so that the expression $\{ \}$ occurring in (11) exceeds, on \mathbf{F}_C^* , a positive lower bound. Hence, on placing

$$(12) \quad \sigma = 2\xi \{C - 3(\xi^2 - \eta^2)^2\}^{\frac{1}{2}}, \quad \tau = 2\eta \{C - 3(\xi^2 - \eta^2)^2\}^{\frac{1}{2}},$$

where $\{ \}^{\frac{1}{2}} > \text{const.} > 0$,

one sees from (11) that the equation of \mathbf{F}_C^* may be written in the form $\dot{\xi}^2 + \dot{\eta}^2 + \sigma^2 + \tau^2 = 8$ of a three-dimensional hypersphere \mathbf{S} in a

four-dimensional $(\xi, \eta, \sigma, \tau)$ -space. However, the correspondence between the points of the hypersurface \mathbf{F}_C^* and the hypersphere \mathbf{S} is not one-to-one. For, as pointed out after (11), the isoenergetic states (ξ, η, ξ, η) and $(-\xi, -\eta, -\xi, -\eta)$ represent one and the same point of \mathbf{F}_C^* . According to (12), this amounts to the identification of the two distinct points $(\xi, \eta, \sigma, \tau), (-\xi, -\eta, -\sigma, -\tau)$ of \mathbf{S} .

Thus, the points of \mathbf{F}_C^* are readily seen to be in one-to-one continuous correspondence with the points of a manifold \mathbf{S}^* which one obtains by identifying the diametrically opposite points of a hypersphere \mathbf{S} . But the manifold \mathbf{S}^* thus defined is identical with the manifold of all lines through the mid-point of \mathbf{S} . Since the latter manifold is the three-dimensional projective space, the proof is complete.

It may be mentioned that, for reasons of continuity, the proof and the result remain unchanged if one replaces (4) by (2), where the limiting case $\mu = 0$ of §300 need not be excluded.

§501. On identifying $(6_1)-(6_2)$, §229 with $(8_1)-(9_1)$, §498, and (21), §232 with (11), §499, one sees that §232 is applicable. Hence, those solutions $\dot{\xi} = \dot{\xi}(\bar{t}), \dot{\eta} = \dot{\eta}(\bar{t}); \xi = \xi(\bar{t}), \eta = \eta(\bar{t})$ of (9_1) , §498 which constitute the three-dimensional manifold \mathbf{F}_C^* may be obtained from a system

$$(13) \quad \dot{\xi} = \Xi(\xi, \eta, \omega; C), \quad \dot{\eta} = H(\xi, \eta, \omega; C), \quad \dot{\omega} = \Omega(\xi, \eta, \omega; C)$$

of three differential equations of the first order for the three variables $\xi, \eta, \omega = \arctan \dot{\eta}/\dot{\xi}$; cf. (24), §232.

In view of (25), §232, this system satisfies the incompressibility condition of §122. Since (13) is obtained from (9_1) by isoenergetic reduction, it is clear from §81 that \mathbf{F}_C^* is an invariant set of (13). And the last remark of §498 implies that all solutions may be considered as unrestricted in the sense of §119; so that \mathbf{F}_C^* is an unrestricted invariant set of (13). Thus, all conditions of §120–§121 are satisfied by (13).

§501 bis. Finally, also the remaining assumption of the Ergodic Theorem (§123–§124) is satisfied in the astronomically significant case of \mathbf{F}_C^* .

In fact, the assumption of the Ergodic Theorem is that the (Euclidean) volume measure of the (ξ, η, ω) -space of (13) is finite. But ω is an angular variable, to be reduced mod 2π ; so that it is sufficient to show that the admissible two-dimensional (ξ, η) -space has a finite

Euclidean (ξ, η) -area. And this condition is satisfied in the astronomically significant case, since in this case §497 shows that the admissible (x, y) -space is practically the small circle $x^2 + y^2 \leq 4C^{-2}$, a circle which, by (7), corresponds to two (ξ, η) -circles, $\xi^2 + \eta^2 \leq 2C^{-1}$.

§502. On replacing the restricted problem of three bodies by the non-planar model of §478, and then repeating the considerations which in §489–§492 led to (4)–(5₁), one readily finds that (4) and (5₁) must be replaced by

$$(14) \quad U = \frac{3}{2}x^2 + \frac{1}{2}z^2 - (x^2 + y^2 + z^2)^{-\frac{1}{2}} \text{ and by } (2_1)-(2_3), \text{ §478.}$$

Periodic Lunar Orbits

§503. The starting point of the modern theory of the Moon is a certain solution $x = x(t)$, $y = y(t)$ of (5₁), §493. This solution, introduced by Hill, represents a motion which is symmetric with respect to each of the coordinate axes $y = 0$, $x = 0$ and is periodic, with a period τ which is an integration constant of (5₁), §493. On choosing the origin of the t -axis in such a way that the Moon is situated on the positive half of the axis of syzygies when $t = 0$, one has $x(0) > 0$, $y(0) = 0$; so that the symmetry requirement is expressed by the four conditions

$$(1) \quad x(-t) = x(t) = -x(t + \frac{1}{2}\tau), \quad -y(-t) = y(t) = -y(t + \frac{1}{2}\tau).$$

This means that the Fourier expansion of the periodic solution, i.e.,

$$x(t) = \sum_{n=0}^{\infty} (\alpha_n \cos \nu n t + \beta_n \sin \nu n t), \quad y(t) = \sum_{n=0}^{\infty} (\gamma_n \cos \nu n t + \delta_n \sin \nu n t),$$

where $\nu = 2\pi/\tau$, is required to be such that, not only $\beta_n = 0$, $\gamma_n = 0$, but also $\alpha_{2n} = 0$, $\delta_{2n} = 0$, for every n . Thus, if $\alpha_{2n+1} = A_n$, $\beta_{2n+1} = B_n$, then

$$(2) \quad x(t) = \sum_{n=0}^{\infty} A_n \cos (2n+1)t/m, \quad y(t) = \sum_{n=0}^{\infty} B_n \sin (2n+1)t/m,$$

$$\text{where } \tau = 2\pi m, \quad (m = 1/\nu).$$

On replacing A_n , B_n by their linear combinations

$$(2 \text{ bis}) \quad 2a_n = A_n + B_n, \quad 2a_{-n-1} = A_n - B_n, \quad \text{where } n = 0, 1, 2, \dots,$$

one can write (2) in the form

$$(3) \quad x(t) = \sum_{k=-\infty}^{+\infty} a_k \cos (2k + 1)t/m, \quad y(t) = \sum_{k=-\infty}^{+\infty} a_k \sin (2k + 1)t/m; \\ m = \tau : 2\pi.$$

Of course, the problem is that of finding solutions $x = x(t)$, $y = y(t)$ which are of the form (1) or (3), where the period $\tau = 2\pi m$ is an integration constant which determines the amplitudes

$$(4) \quad a_k = a_k(m); \quad k = 0, \pm 1, \pm 2, \dots$$

Thus, (3) is an unknown one-parametric family of solutions of (5₁), §493.

§504. The procedure leading to this family of periodic solutions will follow a straightforward program which must be considered as rather bold, since it may be described as follows:

On substituting the Fourier series (3) into the differential equations (5₁), §493, one is led, by comparison of the coefficients of $\cos kt/m$, $\sin kt/m$ for every k , to equations of condition for the unknown coefficients (4); so that there results an infinite system of simultaneous conditions. Let this infinite system of equations be denoted by (S); so that (S) contains all unknown functions (4) and the parameter m , the latter being introduced by the derivatives x' , y' ; x'' , y'' of (3). Since (5₁), §493 is non-linear, so is (S). And (S) is not a recursive system of equations, since each of the equations constituting (S) contains each of the unknowns (4).

Nevertheless, it will be possible to show that the system (S) determines the functions (4) uniquely, at least if the period $2\pi m$, which will be considered as an independent variable, is restricted to a certain range. In order to complete the proof of the existence of the periodic family (3), it will, of course, be necessary to show that the solution (4) of (S) tends, as $k \rightarrow \pm \infty$, to 0 so strongly as to make each of the formal trigonometric series (3) a Fourier series of an (analytic) periodic function of t for every fixed value of m on the m -range under consideration.

Finally, it must turn out that this range of the integration constant m contains the numerical value $m = m_0$ which belongs to the Moon of the Earth (cf. the beginning of §503). Since this value of m is the small number $m_0 = 0.08084 \dots$ (cf. §518 below), the m -range of interest is the immediate vicinity of $m = 0$.

§505. The explicit calculation of (S) is facilitated by replacing the coordinates x, y , the time variable t , and the operator $' = d/dt$ by

$$(5_1) \quad u = x + iy, v = x - iy; \quad (5_2) \quad \zeta = \exp(it/m); \quad (5_3) \quad D = \zeta d/d\zeta,$$

respectively, where $i = +\sqrt{-1}$. First, it is seen from (4), §491 that the Lagrangian equations (5₁), §493 and their energy integral (5₂), §493 become

$$(6_1) \quad u'' + 2iu' - \frac{3}{2}(u + v) = -u/(uv)^{\frac{1}{2}},$$

$$v'' - 2iv' - \frac{3}{2}(u + v) = -v/(uv)^{\frac{1}{2}};$$

$$(6_2) \quad u'v' - \frac{3}{4}(u + v)^2 - 2/(uv)^{\frac{1}{2}} = -C,$$

upon using (5₁), where $uv = x^2 + y^2$, $u'v' = x'^2 + y'^2$. Since $' = d/dt$, it is clear from (5₂)–(5₃), where $m = \text{const.}$ and $i^2 = -1$, that (6₁)–(6₂) may be written as

$$(7_1) \quad D^2u + 2mDu + \frac{3}{2}m^2(u + v) = m^2u/(uv)^{\frac{1}{2}},$$

$$D^2v - 2mDv + \frac{3}{2}m^2(u + v) = m^2v/(uv)^{\frac{1}{2}};$$

$$(7_2) \quad DuDv + \frac{3}{4}m^2(u + v)^2 + 2m^2/(uv)^{\frac{1}{2}} = C.$$

It is understood that D^2 denotes the iterate, DD , of (5₃); so that, for instance,

$$(8) \quad D^2(uv) = uD^2v + 2DuDv + vD^2u; \quad D(Dv - vDu) = uD^2v - vD^2u,$$

since $D(f + g) = Df + Dg$, $D(fg) = fDg + gDf$.

It is clear from (5₁)–(5₂) that the unknown family of periodic solutions (3) may be written as

$$(9) \quad u = \sum_{k=-\infty}^{+\infty} a_k \zeta^{2k+1}, \quad v = \sum_{k=-\infty}^{+\infty} a_{-k-1} \zeta^{2k+1}.$$

Hence, the first point of the program of §504, namely the determination of the system (S) , requires the comparison of the coefficients of the powers of ζ in the pair of equations which one obtains by substituting (9) into (7₁). Although this is quite unmanageable in view of the square root and division signs which occur in (7₁), the difficulty may be removed by expressing $1/(uv)^{\frac{1}{2}}$ in (7₁) as the cube of the polynomial representation which follows from (7₂) for $1/(uv)^{\frac{1}{2}}$.

In fact, on using (8) and (7₂), one readily finds that the two equations of motion (7₁) may be written, for every fixed value of the energy constant C , in the form

$$(10) \quad \begin{aligned} D^2(uv) - DuDv - 2m(uDv - vDu) + \frac{9}{4}m^2(u+v)^2 &= C, \\ D(uDv - vDu) - 2mD(uv) + \frac{3}{2}m^2(u^2 - v^2) &= 0, \end{aligned}$$

which is free of radicals and fractions. Incidentally, the passage from (7₁) to (10) by means of (7₂) is merely the transition from (5₁), §493 to (6), §493 by means of (5₂), §493.

§506. It is clear from (5₂)–(5₃) that the (formal) derivative Df of a Fourier series $f = f(\zeta)$ of the form $\sum \alpha_k \zeta^{2k+1}$ is $\sum (2k+1) \alpha_k \zeta^{2k+1}$; while (5₂) itself implies that if $g = g(\zeta)$ is another Fourier series of the same form, say $\sum \beta_k \zeta^{2k+1}$, then the product fg has the Fourier series $\sum \gamma_k \zeta^{2k}$, where $\gamma_k = \sum \alpha_j \beta_{k-j-1}$ (each of the summation indices k, j runs from $-\infty$ to $+\infty$). On applying these two rules a finite number of times, one sees that substitution of (9) into (10) transforms the two equations (10) into

$$(11) \quad \sum_{k=-\infty}^{+\infty} \mu_k \zeta^{2k} = C, \quad \sum_{k=-\infty}^{+\infty} \nu_k \zeta^{2k} = 0,$$

where μ_k, ν_k are independent of ζ (i.e., of t) and represent polynomials in the parameter m and the infinitely many coefficients (4) together. In view of (11), the system of the equations of condition to be satisfied by the functions (4) is

$$(11 \text{ bis}) \quad \mu_0 = C, \nu_0 = 0; \mu_j = 0 = \nu_j, \text{ where } j = \pm 1, \pm 2, \dots$$

On carrying out explicitly the substitutions mentioned before (11), and then forming suitable combinations of the equations (11 bis) thus obtained, one finds, after straightforward reductions, the following explicit result:* The system of conditions (11 bis) is equivalent to the infinite system which consists, on the one hand, of the two equations represented by

$$(12) \quad \sum_{i=-\infty}^{+\infty} \{ (\overline{2i+1})^2 + 8i + 4m + \frac{9}{2}m^2 \} a_i^2 + \frac{9}{2}m^2 a_i a_{-i-1} \} = C$$

and

$$(13) \quad \left\{ \sum_{i=-\infty}^{+\infty} a_i \right\}^2 \sum_{i=-\infty}^{+\infty} \{ 4i^2 + 4i + 1 + \overline{4i+2} m + 3m^2 \} a_i = m^2,$$

* The details of this elementary calculation may be found in Hill's fundamental memoir.

and, on the other hand, of the simultaneous infinite system

$$(14) \quad \sum_{i=-\infty}^{+\infty} \{ [j, i] a_i a_{i-j} + m^2 [j] a_i a_{-i+j-1} + m^2 (j) a_i a_{-i-j-1} \} = 0;$$

$$j = \pm 1, \pm 2, \dots,$$

where $[j, i]$, $[j]$, (j) are rational functions of the independent variable m , namely

$$(15) \quad [j, i] = -\frac{i}{j} \frac{4(j-1)i + 4j^2 + 4j - 2 - 4(i-j+1)m + m^2}{2(4j^2 - 1) - 4m + m^2},$$

$$(16) \quad [j] = -\frac{3}{16j^2} \frac{4j^2 - 8j - 2 - 4(j+2)m - 9m^2}{2(4j^2 - 1) - 4m + m^2}$$

$$(17) \quad (j) = -\frac{3}{16j^2} \frac{20j^2 - 16j + 2 - 4(5j-2)m + 9m^2}{2(4j^2 - 1) - 4m + m^2},$$

with the understanding that $j = \pm 1, \pm 2, \dots$ but $i = 0, \pm 1, \pm 2, \dots$.

Clearly, the system which in §504 was denoted by (S) consists, on the one hand, of the infinitely many equations (14) whose coefficient functions are given by (15)–(17), and, on the other hand, of the single equation (13). In fact, the rôle of (12) is merely that of supplying the Jacobi constant C as a function $C(m)$ of the parameter m of the family of periodic solutions (3), if the corresponding solution (4) of (S) has already been determined.

§507. It will now be necessary to prove an existence theorem which is applicable to (S) ; cf. §504. In order to formulate this theorem, let a power series F in infinitely many variables z_0, z_1, z_2, \dots be defined, without regard to questions of convergence, as an expression of the type

$$(18) \quad F(z_0, z_1, \dots) = \sum_{n=0}^{\infty} F^{(n)},$$

$$\text{where } F^{(n)} = \sum \dots \sum a_{i_1 \dots i_n}^{(n)} z_{i_1} \dots z_{i_n}$$

is a form of degree n in z_0, z_1, \dots , with n non-negative integral summation indices i_1, \dots, i_n (which need not be distinct), chosen in such a way that the terms of the n -fold series $F^{(n)}$ appear as contracted completely for every n (by this is meant that no monomial in the z_i occurs more than once). For any power series (18), let F^* denote the power series defined by

$$(18 \text{ bis}) \quad F^*(z_0, z_1, \dots) = \sum_{n=0}^{\infty} F^{(n)*},$$

$$\text{where } F^{(n)*} = \sum \dots \sum |a_{i_1 \dots i_n}^{(n)}| z_{i_1} \dots z_{i_n};$$

so that $F^*(z_0, z_1, \dots) \equiv F(z_0, z_1, \dots)$ if and only if all $a \geq 0$.

Let there be given an infinite sequence $F_1(x; y_1, y_2, \dots), \dots, F_k(x; y_1, y_2, \dots), \dots$ of power series (18), where $z_0 = x, z_1 = y_1, \dots, z_k = y_k, \dots$, and suppose that there exist two positive constants and two sequences of positive constants, say $\alpha; \gamma$ and $\beta_1, \dots, \beta_k, \dots; \mu_1, \dots, \mu_k, \dots$, which satisfy the two infinite sequences of inequalities

$$(19_1) \quad F_k^*(\alpha; \beta_1, \beta_2, \dots) \leq \mu_k; \quad (19_2) \quad \beta_k / \mu_k \geq \gamma$$

for every k (the point is that, while $\beta_k > 0$ and $\mu_k > 0$ may depend on k in an arbitrary manner, $\alpha > 0$ and $\gamma > 0$ are supposed to be independent of k).

The existence theorem which will be needed for (S) states that, if (19₁)–(19₂) are satisfied, the infinite implicit system of equations

$$(20) \quad y_k = x F_k(x; y_1, y_2, \dots), \quad (k = 1, 2, \dots),$$

has in a sufficiently small circle about the origin of the complex x -plane, namely at least in the circle

$$(21) \quad |x| < \text{Min}(\alpha, \gamma), \quad (\alpha > 0, \gamma > 0),$$

which is independent of k , exactly one regular analytic solution $y_1 = y_1(x), y_2 = y_2(x), \dots$; and that, for this unique solution $y_k = y_k(x)$ of (20), one has

$$(22_1) \quad |y_k(x)| < \beta_k \text{ for } |x| < \text{Min}(\alpha; \gamma); \quad (22_2) \quad y_k(0) = 0,$$

where $k = 1, 2, \dots$; finally, that $y_k(x)$ is real for real x if all coefficients of each of the power series $F_k(x; y_1, y_2, \dots)$ in infinitely many variables are real.

§508. The proof of this theorem proceeds as follows:

Using the notation defined by (18)–(18 bis), first consider

$$(23) \quad Y_k = x F_k^*(x; Y_1, Y_2, \dots), \quad (k = 1, 2, \dots),$$

instead of (20). The existence theorem stated in §507 must hold for (23) also, since the assumptions (19₁)–(19₂) are the same for (20) as for (23). Thus, one has to prove the existence of exactly one regular

analytic solution $Y_k = Y_k(x) = \sum_{m=0}^{\infty} c_{km} x^m$ of (23) in the circle (21). On denoting the n -th partial sum $\sum_{m=0}^n a_m x^m$ of any (formal) power series $f(x) = \sum_{m=0}^{\infty} a_m x^m$ in x by $[f(x)]_n$, and substituting the infinitely many power series $Y_1(x), Y_2(x), \dots$ (which are not known as yet) into (23), one sees by comparing the coefficients of x^n in $Y_k(x) = xF_k^*(x; Y_1(x), Y_2(x), \dots)$, that if the $Y_k(x)$ exist, their partial sums $[Y_k(x)]_n$ must satisfy

$$\begin{aligned} [Y_k(x)]_n &= [xF_k^*(x; Y_1(x), Y_2(x), \dots)]_n \\ &= x[F_k^*(x; Y_1(x), Y_2(x), \dots)]_{n-1} \\ &= x[F_k^*(x; [Y_1(x)]_{n-1}, [Y_2(x)]_{n-1}, \dots)]_{n-1}. \end{aligned}$$

In other words, the power series $Y_k(x)$ must be chosen so that if $Y_k^n = Y_k^n(x)$ denotes the polynomial $[Y_k(x)]_n$ (of degrees $\leq n$), then

$$(24) \quad Y_k^n(x) = x[F_k^*(x; Y_1^{n-1}(x), Y_2^{n-1}(x), \dots)]_{n-1}.$$

Now, (24) is a recursive system which determines all partial sums $Y_k^n(x)$ of all $Y_k(x)$, as follows:

On placing $x = 0$ in (23), one sees that (22₂) is satisfied by $y_k = Y_k$; so that $Y_k(0) = 0$, i.e., $Y_k^0(x) \equiv 0$. This determines the start of the recursion system (24) for the $Y_k^n(x)$. For instance, application of (24) to $n = 1$ gives $Y_k^1(x) = xF_k^*(0; 0, 0, \dots)$, since $[F_k^*(x; Y_1^0(x), Y_2^0(x), \dots)]_0 = [F_k^*(x; 0, 0, \dots)]_0 = F_k^*(0; 0, 0, \dots)$, by the definition of the operator $[]_n$.

Suppose that, for a fixed $n - 1 \geq 0$ and for every k , one has already determined the polynomials $Y_k^0(x), Y_k^1(x), \dots, Y_k^{n-1}(x)$ in accordance with (24) and in such a way that these polynomials have real, non-negative coefficients and are, in the circle (21), less than β_k in absolute value, where $k = 1, 2, \dots$. These conditions are satisfied for $n - 1 = 0$, since $Y_k^0(x) \equiv 0$. The induction from $n - 1$ to n may easily be carried out. In fact, since $|Y_k^{n-1}(x)| < \beta_k$ in the circle (21), and since the polynomials $Y_k^{n-1}(x)$ have real, non-negative coefficients (as do, by (18 bis), the power series $F_k^*(x; Y_1, Y_2, \dots)$ in infinitely many variables $x; Y_1, Y_2, \dots$), one sees from (19₁) that $F_k^*(x; Y_1^{n-1}(x), Y_2^{n-1}(x), \dots)$ defines in the circle (21) a regular analytic function which may there be reordered into a convergent power series; and that the absolute value of this power series in x cannot exceed μ_k in the circle (21). Furthermore, the coefficients of this power series in x are real, non-negative numbers; so that also the absolute

value of its partial sum $[F_k^*(x; Y_1^{n-1}(x), Y_2^{n-1}(x), \dots)]_{n-1}$ cannot exceed μ_k in the circle (21). Consequently, (24) defines, for every k , a polynomial $Y_k^n(x)$ which has real, non-negative coefficients and satisfies, in the circle (21), the inequality $|Y_k^n(x)| \leq |x| \mu_k$; so that $|Y_k^n(x)| \leq \beta_k$, by (19₂) and (21). This completes the induction from $n-1$ to n .

Since the absolute values of the partial sums $Y_k^n(x)$ of the power series $Y_k(x)$ do not exceed β_k in the circle (21), it is clear that $Y_k(x)$ is convergent, and satisfies the inequality $|Y_k(x)| < \beta_k$, in the circle (21).

This completes the proof of all statements of §507 in case (20) is replaced by (23). But (23) clearly is a majorant system of (20); so that the existence theorem announced in §507 and the statements (22₁)–(22₂) follow for (20) also. Finally, the last remark of §507 is clear from the fact that the partial sums $y_k^n(x)$ of the power series $y_k(x)$ follow in a recursive manner from the analogue to (24):

$$(24 \text{ bis}) \quad y_k^n(x) = x[F_k(y_1^{n-1}(x), y_2^{n-1}(x), \dots)]_{n-1}.$$

§509. In order to apply the existence theorem thus proved to the system (S) of §506, notice first that, by (15), one has $[j, j] \equiv -1$ and $[j, 0] \equiv 0$ for $j = \pm 1, \pm 2, \dots$. Hence, (14) may be written in the form

$$(25) \quad \begin{aligned} a_0 a_j = & 0 + 2m^2[j]a_0 a_{j-1} + 2m^2(j)a_0 a_{-j-1} \sum'_{i=-\infty}^{+\infty} [j, i]a_i a_{i-j} \\ & + m^2[j] \sum''_{i=-\infty}^{+\infty} a_i a_{-i+j-1} + m^2(j) \sum'''_{i=-\infty}^{+\infty} a_i a_{-i-j-1}, \end{aligned}$$

if the marks attached to the summation signs mean that the pairs of summation indices

$$(25 \text{ bis}) \quad i = j, i = 0; \quad i = j-1, i = 0; \quad i = -j-1, i = 0$$

must be omitted in \sum' , \sum'' , \sum''' , respectively. Thus, on dividing (25) by ma_0^2 and placing

$$(26) \quad c_j = \frac{a_j}{ma_0}, \quad \text{where } j = \pm 1, \pm 2, \dots,$$

one sees that (25) is equivalent to

$$\begin{aligned}
 (27) \quad c_j = m \left\{ \sum_{i=-\infty}^{+\infty} [j, i] c_i c_{i-j} + m^2 [j] \sum_{i=-\infty}^{+\infty} c_i c_{-i+j-1} \right. \\
 \left. + m^2(j) \sum_{i=-\infty}^{+\infty} c_i c_{-i-j-1} + 2m[j] c_{j-1} + 2m(j) c_{-j-1} \right\}, \\
 \text{where } j = \pm 1, \pm 2, \dots
 \end{aligned}$$

It will turn out that the existence theorem of §507 is directly applicable to the representation (26)–(27) of (14).

§510. To the foregoing end, let $f^* = f^*(m)$ denote, in accordance with (18)–(18 bis), the power series $|C_0| + |C_1|m + |C_2|m^2 + \dots$ belonging to a function $f = f(m)$ which admits, for small $|m|$, the Taylor expansion $C_0 + C_1m + C_2m^2 + \dots$. Then, for the infinitely many rational functions (15), (16), (17) of m and for a suitable bound B which is independent of i, j and m , one has

$$(28_1) \quad |[j, i]^*| < B(|i/j|^2 + |i/j|) \quad \text{for } |m| \leq 1;$$

$$(28_2) \quad |[j]^*| < B/j^2 \quad \text{and} \quad |(j)^*| < B/j^2 \quad \text{for } |m| \leq 1.$$

In fact, if a function $f(m) = C_0 + C_1m + \dots$ is regular analytic in a circle $|m| < R$, and if $|f(m)| < M = \text{const.}$ in this circle, then $|C_n|$ is known to be less than M/R^n for $n = 0, 1, 2, \dots$. It follows that if $R > 1$, then $|f^*(m)| < MR/(R - 1)$ for $|m| \leq 1$. Hence, on dividing the numerator and the denominator of (15) by $2(4j^2 - 1)$, one sees that, in order to prove (28₁), it is sufficient to assure the existence of an $M > 0$ and an $R > 1$ which have the property that the infinitely many rational functions

$$f_j(m) = \left\{ 1 - \frac{4m - m^2}{2(4j^2 - 1)} \right\}^{-1}, \quad \text{where } j = \pm 1, \pm 2, \dots,$$

are regular analytic and in absolute value less than M in the circle $|m| < R$. But this condition is satisfied by $R = \frac{8}{7}$ and a sufficiently large $M = \text{const.}$, since if $|m| < 1 + \frac{1}{7}$, then

$$\begin{aligned}
 |4m - m^2| &< (1 + \tfrac{1}{7})(4 + 1 + \tfrac{1}{7}) < 6 \leq 2(4j^2 - 1) \\
 &\text{for } j = \pm 1, \pm 2, \dots
 \end{aligned}$$

This proves (28₁); while (28₂) follows from (16)–(17) in the same way as (28₁) follows from (15).

§511. In addition, there will be needed the elementary fact† that there exists a numerical constant, say C , in such a way that‡

$$(28 \text{ bis}) \quad \sum'_{i=-\infty}^{+\infty} i^{-2}(j-i)^{-2} < Cj^{-2} \quad \text{for } j = \pm 1, \pm 2, \dots,$$

where the summation index i runs, for fixed j , through all integers distinct from $i = 0$ and $i = j$. Furthermore, (28 bis) remains valid if one replaces each of the three exponents -2 occurring in it by any negative integer $-3, -4, \dots$, the value of the numerical constant C depending on this integer.§

§512. As a consequence of (28₁)–(28 bis), there exists a sufficiently large constant, A , which has the property that, if the infinitely many variables $m; c_1, c_{-1}, c_2, c_{-2}, \dots$ are restricted to the region

$$(29) \quad |m| \leq 1; \quad |c_j| \leq j^{-4}, \quad (j = \pm 1, \pm 2, \dots),$$

and are otherwise arbitrary, then, for $j = \pm 1, \pm 2, \dots$,

$$(30_1) \quad \sum'_{i=-\infty}^{+\infty} |[j; i]^*| |c_i c_{i-j}| < Aj^{-4};$$

$$(30_2) \quad |[j]^*| \sum''_{i=-\infty}^{+\infty} |m^2 c_i c_{-i+j-1}| < Aj^{-4}; \quad |(j)^*| \sum'''_{i=-\infty}^{+\infty} |m^2 c_i c_{-i-j-1}| < Aj^{-4};$$

$$(30_3) \quad 2|[j]^*| |mc_{j-1}| < Aj^{-4}; \quad 2|(j)^*| |mc_{-j-1}| < Aj^{-4}.$$

In fact, it is clear from (28₁) that, in the region (29), the expression on the left of (30₁) is less than the product of the constant B and of

† This well-known fact is fundamental in Riemann's theory of trigonometric series, as well as in the multiplication theory of these series.

‡ In order to prove the existence of such a constant C , notice that, if j is even, then, on shifting the summation index i by $\frac{1}{2}j$, one can write the sum on the left of (28 bis) in the form

$$\sum'_{i=-\infty}^{+\infty} (i + \tfrac{1}{2}j)^{-2} (\tfrac{1}{2}j - i)^{-2} = \sum'_{i=-\infty}^{+\infty} (i^2 - \tfrac{1}{4}j^2)^{-2}.$$

But the last sum is obviously less than a constant multiple of j^{-2} , since $\sum i^{-2} < +\infty$. This proves (28 bis) for even j . And the proof clearly is the same for odd j .

§ It is seen from the preceding footnote that the proof of this extension of (28 bis) is the same as that of (28 bis) itself.

$$\sum'_{i=-\infty}^{+\infty} |i/j|^2 i^{-4} (i-j)^{-4} + \sum'_{i=-\infty}^{+\infty} |i/j| i^{-4} (i-j)^{-4}.$$

But the sum of these two series is less than

$$j^{-2} \sum'_{i=-\infty}^{+\infty} i^{-2} (i-j)^{-2} + |j|^{-1} \sum'_{i=-\infty}^{+\infty} |i|^{-3} |i-j|^{-3},$$

and so, by (28 bis) and the extension of (28 bis) mentioned at the end of §511, less than $j^{-2} \cdot Cj^{-2} + |j|^{-1} \cdot \bar{C}|j|^{-3} = \text{const. } j^{-4}$. This proves (30₁). And (30₂) follows from (28₂) in the same way as (30₁) followed from (28₁). Finally, (30₃) is clear from (29) and (28₂).

§513. Since the rational functions (15), (16), (17) of m may, by §510, be developed according to non-negative integral powers of m (for small $|m|$), one can write (27) in the form

$$(31) \quad c_j = mG_j(m; c_1, c_{-1}, c_2, c_{-2}, \dots), \quad (j = \pm 1, \pm 2, \dots),$$

where the G_j are power series in the infinitely many variables $m; c_j$ (incidentally, every G is a quadratic polynomial in the infinitely many c_j , but not in m). Since G_j is the coefficient, $\{ \}$, of m on the right of (27), and since (30₁), (30₂), (30₃), where $A = \text{const.}$, hold in the region (29), there exists a sufficiently large $M = \text{Const. } (\leq 5A)$ satisfying

$$(32) \quad G_j^*(1; 1^{-4}, 1^{-4}, 2^{-4}, 2^{-4}, \dots) < Mj^{-4} \text{ for } j = \pm 1, \pm 2, \dots$$

But if one identifies $m; c_1, c_{-1}, \dots$ and G_1, G_{-1}, \dots with $x; y_1, y_2, \dots$ and F_1, F_2, \dots , respectively, then (31) becomes (20). And (32) shows that the conditions (19₁)–(19₂) are satisfied by $\alpha = 1, \gamma = 1/M$. Thus, the circle (21) becomes $|m| < M^{-1}$, where the upper bound M is chosen to be not less than 1.

§514. Consequently, the theorem of §507 assures for (31) the existence of exactly one solution $c_j = c_j(m)$ in the circle $|m| < M^{-1}$, where $c_j(m)$ is regular analytic and in absolute value less than j^{-4} in this circle, while $c_j(0) = 0$; cf. (22₁)–(22₂). But (31) is in virtue of (26) equivalent to (25), i.e., to (14). Hence, the infinitely many conditions (14) for the infinitely many unknown functions (4) of m define the ratios a_j/a_0 in a circle $|m| < M^{-1}$ as uniquely determined regular analytic functions in such a way that

$$(33) \quad a_j/a_0 = m^2 P_j(m) \quad \text{and} \quad |m P_j(m)| < j^{-4} \quad \text{for} \quad |m| < M^{-1},$$

$$(j = \pm 1, \pm 2, \dots),$$

where $P_j(m)$ is a regular power series with real coefficients.

On calculating the first partial sums of these power series by means of the recursion formula (24 bis), one finds from (25) and (15)–(17) that

$$\frac{a_1}{a_0} = \frac{3}{16} m^2 + \frac{1}{2} m^3 + \frac{7}{12} m^4 + \frac{11}{36} m^5 - \frac{30749}{2^{12} \cdot 3^3} m^6 - \dots,$$

$$\frac{a_{-1}}{a_0} = -\frac{19}{16} m^2 - \frac{5}{3} m^3 - \frac{43}{36} m^4 - \frac{14}{27} m^5 - \frac{7381}{2^{10} \cdot 3^4} m^6 + \dots,$$

$$(33 \text{ bis}) \quad \frac{a_2}{a_0} = \frac{25}{256} m^4 + \frac{803}{1920} m^5 + \frac{6109}{2^5 \cdot 3^2 \cdot 5^2} m^6 + \dots,$$

$$\frac{a_{-2}}{a_0} = 0 \cdot m^4 + \frac{23}{640} m^5 + \frac{299}{2^5 \cdot 3 \cdot 5^2} m^6 + \dots,$$

$$\frac{a_3}{a_0} = \frac{833}{2^{12} \cdot 3} m^6 + \dots, \quad \frac{a_{-3}}{a_0} = \frac{1}{192} m^6 + \dots, \quad \frac{a_4}{a_0} = \dots$$

That (33) supplies only the ratios of the unknown functions (4), is due to the fact that, thus far, only the infinitely many quadratic homogeneous conditions (14) were used, whereas the system (S) of §504 consists of (14) and of the inhomogeneous condition (13) together (§506). Correspondingly, (13) can now be used to determine $a_0 = a_0(m)$. To this end, it is sufficient to write (13) in the form

$$(34) \quad a_0 = m^{\frac{1}{2}} \left(\sum_{i=-\infty}^{+\infty} a_i/a_0 \right)^{-\frac{1}{2}} \left(\sum_{i=-\infty}^{+\infty} \{4i^2 + 4i + 1 + \overline{4i + 2} m + 3m^2\} a_i/a_0 \right)^{-\frac{1}{2}},$$

and to observe that the expression on the right of (34) becomes a known function of m in virtue of (33). In particular, on using the approximation (33 bis) to (33), one obtains

$$(34 \text{ bis}) \quad a_0 = m^{\frac{1}{2}} (1 - \frac{2}{3} m + \frac{7}{18} m^2 - \frac{4}{81} m^3 + \dots).$$

Finally, all the unknown functions (4) of m follow from (33)–(34), or, approximately, from (33 bis)–(34 bis).

§514 bis. It is clear from §506 that the remaining relation, (12), expresses merely the fact that the series (3) satisfy, at least formally, not only the equations of motion, represented by (12)–(13), but the energy integral also. Correspondingly, on substituting on the left of (12) the functions (4) which now are supplied by (33)–(34), one obtains the energy constant as a function of the period (cf. §100):

$$(35) \quad C = C(m) = m^{-3}(1 + \frac{8}{3}m + \frac{7}{18}m^2 - \frac{140}{81}m^3 - \dots),$$

by (33 bis)–(34 bis).

§515. The existence of the family of periodic solutions (3) is now established for all, positive and negative, values of the parameter m which are sufficiently small in absolute value.

In fact, (33)–(35) were established for sufficiently small $|m|$. And the $a_j = a_j(m)$ were, in §506–§514, determined in such a way that the series (3) formally satisfy, for fixed m , the Lagrangian equations $x'' - 2y' = U_x$, $y'' + 2x' = U_y$. But the estimate (33) of the a_j assures that the trigonometrical series (3) not only are Fourier series but are Fourier series of functions $x = x(t)$, $y = y(t)$ with continuous second derivatives $x''(t)$, $y''(t)$ which may be obtained by formal differentiation of the series (3). In fact, the j -th Fourier coefficient of $x''(t)$ or $y''(t)$ is, in view of (33), majorized by $j^2 \cdot j^{-4} = j^{-2}$, while $\sum j^{-2} < +\infty$. Inasmuch as the series (3) satisfy the Lagrangian equations formally, it is now clear from the uniqueness theorem of Fourier series, that (3) represents, for fixed m , a solution of the Lagrangian equations.

Lunar Theory

§516. The coefficients of (3) tend with $1/k$ much more strongly to 0 than in the order used in §515. In fact, the equations of motion are regular analytic in x, y (if one excludes the origin $x = 0, y = 0$); so that their solutions $x = x(t)$, $y = y(t)$ are regular analytic in t (if one excludes collisions). But it is known (O. Hölder) that a periodic trigonometrical series is Fourier series of a periodic, regular analytic function if and only if the coefficients tend to 0 as strongly as the terms of a convergent geometrical progression. Hence, in (3) one has $|a_k(m)| < q^{|k|}$, where $q = q(m) > 0$ is less than 1 and independent of k .

But more than this is true. In fact, on calculating from (31), i.e.,

from (27) and (26), the power series (33) by using the recursion formula (24 bis), one readily verifies from (14)–(17) by complete induction that the power series a_j/a_0 in m vanishes at $m = 0$ in at least the $2|j|$ -th order (this fact has already been indicated by the approximations (33 bis), which are calculated precisely in this manner). But the principle of the maximum for regular analytic functions is known to imply the lemma (H. A. Schwarz) according to which a power series of the form $f(z) = \alpha_n z^n + \alpha_{n+1} z^{n+1} + \dots$, where $n \geq 1$, cannot be convergent and in absolute value less than a constant μ in a circle $|z| < \rho$, unless $|f(z)| < |z|^n \mu / \rho^n$ in this circle. It follows, therefore, from (33) that there exists, for every positive number K which is less than M^{-1} , a positive number L such that $|a_j(m)/a_0(m)| < Lm^{2|j|}$ holds for $|m| < K$ and $j = \pm 1, \pm 2, \dots$. This puts into explicit evidence the existence of a $q = q(m) < 1$.

§517. It is now easy to find the dynamical significance of the periodic family (4) for small values of the integration constant m (≥ 0). In fact, on neglecting in (4) the coefficients $a_j = a_j(m)$, $j = \pm 1, \pm 2, \dots$ which are of a higher order in m than is $a_0 = a_0(m)$, one obtains

$$(36) \quad \begin{aligned} x &= a_0(m) \cos(t/m), & y &= a_0(m) \sin(t/m), \\ \text{where } a_0(m) &= m^{\frac{1}{3}} - \dots; & C(m) &= m^{-\frac{1}{3}} + \dots, \end{aligned}$$

by (34 bis), (35). This approximation (36) defines the synodical path $x = x(t)$, $y = y(t)$ of the Moon as a uniform circular motion of radius a_0 about the position $(x, y) = (0, 0)$ of the Earth, the period $2\pi m$ being reckoned as positive or negative according as this synodical motion is direct or retrograde. And the assumption of (36) is that the integration constant m , i.e., the approximate “radius” $a_0 = m^{\frac{1}{3}} - \dots$, is very small. Accordingly, the influence of the Sun becomes negligible, and so the model is practically a problem of two bodies (Earth-Moon), considered, as in §300, from a synodical coordinate system; so that §307 is applicable in the circular case. But the parameter m , when defined by (13)–(14₁), §306–§307, is the ratio of the synodical period and of 2π ; while (14₂)–(14₃), §307, show that the radius and the Jacobi constant become $m^{\frac{1}{3}} - \dots$ and $m^{-\frac{1}{3}} + \dots$. Since this agrees with (36), it follows that, if $|m|$ is very small, the parameter m of the periodic family (3) may be identified with (14₁), §307.

Accordingly, one can interpret the periodic family (3) as follows:

If the Sun did not disturb the system Earth-Moon, (4) would be identical with the circular family considered in §307; and what has been proved in §503–§516 is to the effect that the presence of the Sun perturbs the system Earth-Moon in such a way as to preserve, at least for sufficiently small $|m|$, the existence of a periodic family, this family being precisely (4) and, for very small $|m|$, approximately (36).

§518. In particular, (13)–(14₁), §306–§307 show that if the perturbations exerted by the Sun were negligible, the value $2\pi m_0$ of the integration constant $2\pi m$ which belongs to the Moon of the Earth could be defined as the synodical period (month) of the lunar path, which then is exactly circular. The value of m_0 mentioned at the end of §504 is somewhat less than 1:12 and corresponds to the actual empirical value of the synodical period $2\pi m_0$.

On carrying out the estimates of §510–512 explicitly, one finds that this m_0 is “sufficiently small” from the point of view of the existence theorem, i.e., that $m_0 (> 0)$ is less than a value of the bound M^{-1} occurring in (33). Since the proof requires only straightforward numerical calculations, it will not be reproduced here.

§518 bis. In view of the possible complex singularities of the power series (33)–(33 bis), it may be expected that the range of convergence of these expansions will be enlarged for positive m , if one subjects the expansion parameter m to what in the theory of divergent series is called an Euler transformation. Such a transformation of m is a linear substitution of the form $m^* = m/(1 - \kappa m)$, where κ is a positive number, to be chosen at convenience. Since the (complex) singularities of the rational coefficient function (28 bis) are closest to the origin $m = 0$ when $j^2 = 1$, Hill found it convenient to take care first of all of the denominator $6 - 4m + m^2$ in the coefficient functions (14)–(17). To this end, $\kappa = \frac{1}{3}$ appeared to be a favorable choice of the constant which determines the Euler transformation.

§519. However, such explicit summation methods cannot help, if one is interested in the periodic solution (3)–(4) in a case in which the integration constant $|m|$ is not small enough. In such a case, recourse has to be made to mechanical quadratures. Then it turns out that, while the curve of zero velocity belonging to the periodic orbit surrounds the latter if $|m|$ is sufficiently small, the periodic orbit reaches its curve of zero velocity when m tends increasingly to

the positive value $0.56096 \dots$ (which is much larger than $1:12$); and that the cusp, which is acquired (by §238) for this particular m , appears on the y -axis. Finally, when the integration constant m of (3)–(4) passes increasingly through the value which belongs to this cuspidal periodic orbit, the cusp develops, in accordance with §240, into a small loop which seems to increase rapidly when m increases further. Unfortunately, it has thus far been impossible to carry out the mechanical quadratures to a stage of the periodic family (3)–(4) sufficiently advanced to indicate the ultimate fate of this family, when this process of analytic prolongation (in m) is continued indefinitely.

§519 bis. All that is certain is that the family is subjected to what E. Strömgren has empirically formulated, on the basis of his numerical material, as the Principle of Natural Termination. This general principle, for which to-day a rigorous mathematical proof is available, does not lie within the scope of this book.

§520. Consider the solution (3)–(4) of

$$(37) \quad x'' - 2y' = U_x(x, y), \quad y'' + 2x' = U_y(x, y) \quad (\text{cf. §493})$$

for a fixed m . According to §234, the corresponding Jacobi equations are

$$(38) \quad \begin{aligned} \xi'' - 2\eta' &= U_{xx}(t; m)\xi + U_{xy}(t; m)\eta; \\ \eta'' + 2\xi' &= U_{xy}(t; m)\xi + U_{yy}(t; m)\eta, \end{aligned}$$

where $U_{xx}(t; m), \dots$ denote the functions which one obtains by substituting (3)–(4) into $U_{xx}(x, y), \dots$; so that the coefficients of (38) are, for fixed m , given periodic functions of t , with $2\pi m$ as period. Hence, the linear system (38) determines four multipliers s_1, s_2, s_3, s_4 (§143) which, by §149, may be grouped into two pairs of the form $(1, 1), (s, 1/s)$. It is understood that the multiplier s is a function $s(m)$ of m . It will be assumed that the fixed value of m under consideration is such that $|s| = 1$ but $s \neq \pm 1$. Numerical calculations show that these conditions are satisfied in an m -range which contains the small value $m_0 = 0.0808 \dots$ belonging to the Moon of the Earth. It will be assumed that $|m|$ is sufficiently small.

§520 bis. On comparing the two-fold symmetry (§503) of the periodic solution (3) with the results of §144, one readily sees that (38)

has, corresponding to the pair of multipliers $(s, 1/s)$, two linearly independent solutions of the form

$$(39) \quad \begin{aligned} \xi &= \epsilon \sum_{k=-\infty}^{+\infty} \alpha_k \cos \{ (2k + 1 + \lambda)t/m + \delta \}, \\ \eta &= \epsilon \sum_{k=-\infty}^{+\infty} \beta_k \sin \{ (2k + 1 + \lambda)t/m + \delta \}, \end{aligned}$$

where $\delta, \epsilon (\neq 0)$ is an arbitrary pair of real integration constants, while the real data

$$(39 \text{ bis}) \quad \lambda = \lambda(m) \text{ and } \alpha_k = \alpha_k(m), \beta_k = \beta_k(m), (k = 0, \pm 1, \dots),$$

are uniquely determined by m . In particular, $\lambda = \lambda(m)$ represents the characteristic exponent belonging to $s = s(m)$; cf. §143–§144, where the normalization of the characteristic exponent is different (the period being thought of as submerged into λ).

Since $s \neq \pm 1$, it is clear that the periodic solution $\xi = x', \eta = y'$ of (38), which is supplied by §148, is linearly independent of the two almost periodic (and, if $\lambda = \lambda(m)$ is rational, periodic) solutions represented by (39). Finally, application of the rule of §149 to the family (3)–(4) supplies the fourth solution of (38), at least if m does not belong to a set of isolated values. Since this fourth solution of (38) contains, by §149, a secular term, it is readily seen from the formulae of §235–§237 bis that the three linearly independent solutions of (38) which correspond to isoenergetic displacements are represented by (39) and the trivial solution $(\xi, \eta) = \text{const.}(x'(t), y'(t))$.

In what follows, only the non-trivial isoenergetic displacements (39) will be considered. It will be assumed that the integration constants δ, ϵ occurring in (39) have fixed values, and that $\epsilon \neq 0$.

§521. First, it is clear from a classical continuity theorem, concerning systems of ordinary linear differential equations, that the functions (39 bis) depend on m continuously.

Since $\lambda = \lambda(m)$ is readily found to be dependent on m , it follows that the almost periodic functions (39) become periodic for a dense but enumerable set of values of m ; the period of (39) for such m being the longer the higher is the commensurability $\lambda(m):1$.

As far as the coefficient functions $\alpha_k(m), \beta_k(m)$ are concerned, one can show that their behavior for large $|k|$ and fixed small $|m|$ is about the same as the behavior of the $a_k(m)$, described in §516. In

particular, the limiting values $\alpha_k^0 = \lim \alpha_k(m)$, $\beta_k^0 = \lim \beta_k(m)$ corresponding to $m \rightarrow 0$ vanish unless $|2k + 1| = 1$. On the other hand, if $|2k + 1| = 1$, i.e., if those terms of (39) are considered which belong to $k = 0$ and $k = -1$, explicit calculations show that

$$(40) \quad \alpha_0^0 \neq 0, \beta_0^0 = \alpha_0^0; \quad \alpha_{-1}^0 = -3\alpha_0^0, \beta_{-1}^0 = 3\beta_0^0$$

$$(\text{while } \alpha_0^k = 0 = \beta_0^k, \text{ if } 2k + 1 \neq \pm 1).$$

Incidentally, one can verify (40) by comparison with the formulae belonging to Keplerian circular motion (cf. §517).

§522. On placing $u = t/m$, then letting $m \rightarrow 0$ in (39)–(39 bis), and omitting the multiplicative integration constant $\epsilon \neq 0$ (or, rather, $2\epsilon\alpha_0^0 \neq 0$), one finds from (40) after an easy reduction that, if λ^0 denotes the limiting value $(\lambda)_{m=0}$ of the characteristic exponent $\lambda = \lambda(m)$, then

$$(\xi)_{m=0} = -\cos u \cos (\lambda^0 u + \delta) - 2 \sin u \sin (\lambda^0 u + \delta),$$

$$(\eta)_{m=0} = -\sin u \cos (\lambda^0 u + \delta) + 2 \cos u \sin (\lambda^0 u + \delta).$$

Hence, $(\xi^2 + \eta^2)_{m=0} = \cos^2 (\lambda^0 u + \delta) + 4 \sin^2 (\lambda^0 u + \delta)$; so that the continuous function $(\xi^2 + \eta^2)_{m=0}$ of u is positive and periodic and has, therefore, a positive lower bound for $-\infty < t < +\infty$. Thus, it is clear for reasons of continuity that, if $|m|$ is sufficiently small, the almost periodic (and, if $\lambda = \lambda(m)$ is rational, periodic) functions (39) are such that $\xi^2 + \eta^2 > \text{const.} > 0$ for $-\infty < t < +\infty$; the value of const. depending on the integration constants $\delta, \epsilon (\neq 0)$ and on m .

Consequently, if $|m|$ is sufficiently small, the theorem mentioned in §484 is applicable to the almost periodic function $\xi(t) + i\eta(t)$ defined by (39). This means that the angular function $\tilde{\omega} = \tilde{\omega}(t)$ defined by $\xi = (\xi^2 + \eta^2)^{\frac{1}{2}} \cos \tilde{\omega}$, $\eta = (\xi^2 + \eta^2)^{\frac{1}{2}} \sin \tilde{\omega}$ admits a decomposition $\tilde{\omega}(t) = \mu t + \chi(t)$ into a secular term μt and an almost periodic remainder term $\chi(t)$, where the mean motion $\mu = \mu(m)$ and the frequencies of $\chi(t)$ are contained in the integral modul of the frequencies of (39).

§523. In particular, the determination of the mean motion $\mu = \mu(m)$ depends, for fixed m , on the determination of the characteristic exponent $\lambda = \lambda(m)$ and of the values of the integers j, l by means of which μ is representable in the form $\mu = (j\lambda(m) + l)/m$.

On determining the actual values of the integers j, l by letting $m \rightarrow 0$, Hill was thus able to reduce the determination of the principal term, μ , in the mean motion of the lunar perigee to that of the characteristic exponent λ . Actually, comparison of §520 with §235–§237 bis shows that the characteristic exponent λ of (39) may be determined also from the equation of isoenergetic normal displacements, an equation of the form $n'' + \kappa(t)n = 0$, where $\kappa(t)$ is a given periodic function of t .

§524. In connection with the foregoing considerations, Hill was led to the method of infinite determinants; while Adams (who calculated the existence of Neptune somewhat before Leverrier), arrived at this method (before Hill) in connection with the inclination problem (8), §481.

This classical method of infinite determinants, as mathematically legalized by Poincaré, is to-day presented, for instance, by the majority of introductory text-books on linear differential equations in the complex domain. Thus, it may suffice to say that this method serves the purpose of furnishing a convenient way for the actual calculation of the characteristic exponent and of the corresponding solutions (10₁), §144; whereas the considerations of §140–§144 assure only the existence of the characteristic exponents and of the corresponding solutions, without supplying a suitable method for their computation.

Notice that, although the method is again that of infinitely many variables, the differential equation and the equations of condition are now linear, instead of being, as in §506–§514, non-linear.

§525. On adding (39) to (3), one obtains two almost periodic functions $x + \epsilon\xi, y + \epsilon\eta$ of t which, in view of §86, represent an approximate solution of (37). It is natural to ask whether or not one can extend this approximate solution of (37) into two almost periodic double series which represent actual solutions of (37). This question, to which the practice of the lunar theory tacitly assumes an affirmative answer, represents a rather difficult mathematical problem which has thus far escaped all the analytical and topological efforts devoted to it.

It is clear that the situation depends very much on whether or not the commensurability condition, mentioned in §521, is satisfied.

(I). The treatment of the first case is relatively easy, though involved enough to lie beyond the scope of this book. In fact, the

treatment of this case of a commensurable characteristic exponent depends on the general theory of periodic solutions of dynamical systems with two degrees of freedom.

(II). The second case, that of an incommensurable characteristic exponent, represents the fundamental difficulty of Celestial Mechanics.

The treatment of this case leads, at least formally, to an infinite process of iterated quadratures. It will now be shown that every single quadrature in this process leads to questions of Diophantine sensitivity and intricacy.

§526. Each of the quadratures in question is one which may be illustrated by that assigned to an $f = f(t)$ by

$$(41) \quad f'(t) = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \mu^{n+m} \cos (n - \alpha m)t,$$

where μ and α are given positive constants, $\mu < 1$, while α has an irrational value; so that the series (41) defines the function $f'(t)$ for $-\infty < t < +\infty$ as an almost periodic, but not periodic, function. It is understood that almost periodicity is meant in the sense of H. Bohr.

If the initial value $f(0)$ is assigned to be 0, then, from (41),

$$(42) \quad f(t) = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \frac{\mu^{n+m}}{n - \alpha m} \sin (n - \alpha m)t.$$

In fact, since $0 < \mu < 1$, the double series (41) is (absolutely and) uniformly convergent for $-\infty < t < +\infty$; so that term-by-term integration is admissible. For the same reason, the double series (42) is absolutely and uniformly convergent for $-T \leq t \leq T$, where $T > 0$ is arbitrarily large but fixed. While this implies that (42) is absolutely convergent for $-\infty < t < +\infty$, it does not follow that (42) is uniformly convergent for $-\infty < t < +\infty$.

§527. It will turn out that (42) is not, in general, a bounded function of t for $-\infty < t < +\infty$. This fact is of historical interest, since the founders of Celestial Mechanics tacitly assumed that questions of stability may be answered in the affirmative by proving the coordinates involved representable by trigonometrical series of the type (42).

§527 bis. What seems to be true is precisely the opposite of this tacit assumption (now disproved), if stability is meant in the sense of §131. In other words, the appearance of the “small divisors” $n - \alpha m$ in (42) might be a formal manifestation of the general situation mentioned in §127 and §131 (cf. also the footnote to §123).

It may be observed in this connection that a formal treatment of the problems considered in §487 and §522 would automatically lead to small divisors, which turned out to be harmless only because it was possible to replace a formal treatment by a suitable application of the general theorem of §484.

§528. In order to discuss the question of the boundedness of f (for $-\infty < t < +\infty$), notice first that the derivative (41) of (42) is almost periodic. It follows, therefore, from a standard theorem on almost periodic functions (P. Bohl), that $f(t)$ is bounded if and only if it is almost periodic.

On the other hand, a necessary (but in itself not sufficient) condition for the almost periodicity of (42) is expressed by the convergence of the square sum of the amplitudes, i.e., by

$$(43) \quad \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \mu^{2(n+m)} / (n - \alpha m)^2 < +\infty.$$

Furthermore, if (43) is satisfied for a fixed $\mu = \mu_0 > 0$ and for some α , then, not only is (43) satisfied for every positive $\mu < \mu_0$ and for the same α , but also

$$(43 \text{ bis}) \quad \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \mu^{n+m} / |n - \alpha m| < +\infty$$

holds for $0 < \mu < \mu_0$ and for the same α . In fact, if (43) holds for a $\mu = \mu_0 > 0$, then, on choosing any positive $\theta < 1$ and placing $\mu = \theta\mu_0$, one readily sees from the inequality $(\sum |a_i b_i|)^2 \leq (\sum a_i^2)(\sum b_i^2)$ that (43 bis) is satisfied. Conversely, (43 bis) is sufficient for (43), since if $\sum |c_i| < +\infty$, then also $\sum c_i^2 < +\infty$.

But (43 bis) implies that the series (43) is uniformly convergent for $-\infty < t < +\infty$ and represents, therefore, an almost periodic function. Consequently, on considering, for every fixed α , the least upper bound, say $\Lambda = \Lambda(\alpha)$, of all those non-negative numbers μ which satisfy either, hence both, of the conditions (43), (43 bis), one arrives at the following result:

There exists for every positive irrational number α a unique non-negative number $\Lambda = \Lambda(\alpha)$ in such a way that

(i) if the value of the function Λ (which is undefined for rational $\alpha > 0$) at a given α is 0, then the function (42) of t is, for this α and for an arbitrarily small $\mu > 0$, neither bounded nor almost periodic;

(ii) if, on the other hand, α is such that $\Lambda = \Lambda(\alpha)$ is not 0, then the function (42) is, for this α and for a positive μ , bounded and almost periodic or unbounded and not almost periodic according as $\mu < \Lambda(\alpha)$ or $\mu > \Lambda(\alpha)$; the limiting case $\mu = \Lambda(\alpha) > 0$ remaining doubtful.

§528 bis. Needless to say, $0 \leq \Lambda(\alpha) \leq 1$. In fact, since $\sum \sum \mu^{n+m}$ is convergent only for $\mu < 1$, it is clear from (43 bis) that if $\Lambda(\alpha) > 1$ for an α , then $|n - m\alpha| \rightarrow \infty$ as $n + m \rightarrow +\infty$. But this is impossible for every fixed α , since it is known that there exist for every irrational $\alpha > 0$ infinitely many pairs of positive integers n_k, m_k such that

$$|\alpha - n_k/m_k| < 1/m_k^2 \text{ for } k = 1, 2, \dots, \text{ where } m_k \rightarrow \infty \text{ as } k \rightarrow \infty$$

(cf. the proof of (ii), §125).

§529. The result of §528 reduces the problem to the investigation of the function $\Lambda(\alpha)$ which is defined for all irrational $\alpha > 0$ as the least upper bound of those $\mu \geq 0$ which satisfy (43 bis). It turns out that $\Lambda(\alpha)$ is a rather discontinuous function of α . In fact, while $\Lambda(\alpha) = 1$ holds for a dense set of α -values, not only does $\Lambda(\alpha) = 1$ fail to hold for some α but one actually has $\Lambda(\alpha) = 0$ on a dense set of α -values. This, and much more, may be proved as follows:

On the one hand, those α for which $\Lambda(\alpha)$ is 0 form a set which is on any α -interval of the second category in the sense of Baire, and is, therefore, such as to contain a non-enumerable set of points on an α -interval of arbitrarily small length and of arbitrary position on the α -axis. This follows by observing that (43 bis) is a particular case of the series which in the theory of real functions are called Borel series.*

* It is interesting that the astronomer Bruns was led to the series (43 bis), and to a quite precise study of its pathological behavior, much earlier (1884) than the general theory of Borel series was developed by the mathematicians. Similarly, the proof of Bruns for the non-enumerability of those α for which $\Lambda(\alpha)$ is 0 seems to be one of the earliest instances of what to-day is called the argument of Baire.

On the other hand, $\Lambda(\alpha) = 1$ almost everywhere. In other words, the set of those α for which $0 \leq \Lambda(\alpha) < 1$ has the Lebesgue measure 0. This result* is a direct consequence of a sharper theorem concerning Diophantine approximations. In fact, it is known that there exist, not only for every algebraic irrational number α , but for almost every irrational number α , two positive numbers $c = c(\alpha)$, $C = C(\alpha)$ such that $|\alpha - n/m| > C/m^c$ holds for arbitrary integers n, m . And the existence of such a pair $c = c(\alpha)$, $C = C(\alpha)$ implies that (43 bis) is satisfied for every $\mu < 1$, i.e., that $\Lambda(\alpha) = 1$.

* Expressed in terms of "geometrical probabilities" by the astronomer Gylden much earlier (1888) than the mathematicians developed the theory of measure.

HISTORICAL NOTES AND REFERENCES

HISTORICAL NOTES AND REFERENCES

The following pages supply references and additions to the successive sections of the text, and contain a few historical remarks of possible interest.

The content, though not the presentation, of the topics treated in Chap. I and Chap. II is so classical that it did not appear to be feasible to give references to the same extent as in the case of the later chapters.

Chapter I

The following monographs are fundamental (also for Chapter II and Chapter III): C. G. J. Jacobi, *Vorlesungen über Dynamik* (1866 [1842–1843]; later (1884) reprinted as *Supplementband* of his *Werke*); H. Poincaré, *Les Méthodes Nouvelles de la Mécanique Céleste*, 1 (1892), 2 (1893), 3 (1899); G. D. Birkhoff, *Dynamical Systems* (1927); T. Levi-Civita–U. Amaldi, *Lezioni di Meccanica Razionale* (three vols., without year).

References to the classical literature of the theory of canonical systems may be found in Cayley's report (1857; *Papers* 3, 156–204) and in some of the standard text-books (in particular, in E. T. Whittaker's *Treatise on the Analytical Dynamics of Particles and Rigid Bodies*, 3rd ed., 1927). It would be rather desirable to make a detailed critical study of the historical development. In fact, the traditional references to the origin of the fundamental mathematical notions in analytical dynamics are almost always incorrect.

For instance, the "Legendre transformation" (§5–§7) is due not to Legendre but to Euler, if not to Leibniz (cf. P. Stäckel, *Bibl. Math.* (3) 1 (1900), 517). Similarly, the introduction of the momenta instead of the velocities occurs in the writings of Lagrange and Poisson, so that the name "Hamiltonian equations" is not justified. In addition, the "Hamilton-Jacobi theory" is only a particular case of Cauchy's theory of characteristics, which is of an older date.

In these circumstances, an attempt has been made to keep down to a minimum the number of definitions associated with a name. Nevertheless, the terminology applied often turns out to be inconsistent from the historical point of view (for instance, the "Lagrangian" derivatives could be called "Eulerian," or, at least, "Euler-Lagrangian").

While instances of §11–§13 are implied by a classical deduction of the ten conservation integrals (cf. the references to §315–§320 below), the full generality of the formalism involved became manifest, via the theories of Lie, only in connection with the conservation principles in the general theory of relativity. For references cf. E. Hölder, *Math. Ztschr.* 31 (1930), 198–201, 230–231.

The presentation of the theory of canonical transformations in the text follows the approach used in the linear case (§57–§64) by A. Wintner, *Ann. di Mat.* (4) 13 (1934), 105–112, and subsequently transferred to the general case (§26–§38) by E. R. van Kampen and A. Wintner, *Amer. Journ. of Math.* 58 (1936), 851–863; cf. also E. R. van Kampen and A. Wintner, *Trans. Amer. Math. Soc.* 44 (1938), 168–195.

The unique polar factorization (§59) of non-singular matrices is contained, at least implicitly, in a paper of L. Autonne (*Palermo Rend.* 16 (1902), 123–125; cf., in fact, A. Wintner, loc. cit., footnote ¹¹). As to the singular case, cf. J. Williamson, *Bull. Amer. Math. Soc.* 41 (1935), 118–123; also 45 (1939), 920–922.

Chapter II

In the same way as before, the following references concern only recent developments, and investigations which are not covered in the works mentioned at the beginning. Concerning the historical development of the topics up to the middle of the 19th century, cf. Cayley's report.

As to §100, cf. the papers (1870) of R. Clausius, L. Boltzmann and C. Szily, reviewed by Boltzmann in the *Fortschr. d. Physik* 26 (1875), 453–460; also E. Betti, *Ann. di Mat.* (2) 8 (1877), 301–311; P. Bohl, *Ztschr. für Math.* 35 (1890), 188–191; G. Herglotz, *Seeliger-Festschrift*, 197–199 (1924).

Poincaré's proof (*Acta Math.* 13 (1890), 67–73) of his recurrence theorem (§123 bis) is perfectly correct, although he does not make explicit reference to the notion of a zero set (the notion of a Lebesgue measure being of a later date). The modernized formulation of Poincaré's theorem was pointed out by E. B. Van Vleck (*Bull. Amer. Math. Soc.* 21 (1915), 335). The ergodic theorem (§123) was proved by G. D. Birkhoff (*Proc. Nat. Acad. Wash.* 17 (1931), 656–666, 650–655; cf. *Bull. Amer. Math. Soc.* 38 (1932), 361–379). The notion of metrical transitivity (§124 bis) was introduced by him and P. A. Smith (*Journ. de Math.* (9) 7 (1928), 360–368). Concerning the

distributional formulation of the ergodic theorem (§123–§124), cf. A. Wintner, *Proc. Nat. Acad. Wash.* 18 (1932), 248–251; P. Hartman and A. Wintner, *Amer. Journ. of Math.* 61 (1939), 977–984. As to a corresponding formulation of the classical circle problem of Poincaré and Denjoy, cf. D. C. Lewis, Jr. and A. Wintner, *Amer. Journ. of Math.* 56 (1934), 407–410. The notion of distribution stability (footnote to §123) was proposed by A. Wintner, *Nature* 145 (1940), 225–226. Concerning the results mentioned in the footnote to §124, cf. J. Hadamard, *Journ. de Math.* (5) 3 (1897), 382–383 and G. D. Birkhoff, *Bull. Soc. Math. de France* 40 (1912), 305–323.

Concerning systems of known transitivity, cf. the report of G. A. Hedlund, *Bull. Amer. Math. Soc.* 45 (1939), 241–260. The great difficulties of all problems of this type can be seen even from the elementary case considered by R. H. Fox and R. B. Kershner, *Duke Math. Journ.* 2 (1936), 147–150. The planar limiting case of the ellipsoid problem (cf. §202 bis) was pointed out by W. Wirtinger, *Jahresber. d. D. M. V.* 9 (1900), 130–131.

As to §125–§130, cf. T. Levi-Civita, *Prace Mat.-Fiz.* 17 (1904), 35–38; *Atti del Congr. Intern. Fis.* 1927, 1–39; *Abh. Math. Sem. Hamburg* 6 (1928), 326–366.

The stability criterion of §132–§133 is due, in the main, to Poincaré and Birkhoff (cf. their works referred to at the beginning of the references to Chap. I). The criterion as given in the text does not assume the restriction that the point-transformation be volume-preserving. Correspondingly, it is assumed that stability is referred to both past and future.

The example of §135 was given, in a slightly different form, by P. Painlevé (*Comptes Rendus* 138 (1904), 1555–1557), that of §136 bis by T. M. Cherry (*Trans. Cambr. Phil. Soc.* 23 (1925), 199–200). To the footnote of §134 cf. H. Bruns, *Berl. Sitzber.* 1890, 543–545 and (concerning F. Minding) P. Stäckel, *Jahresber. d. D. M. V.* 14 (1905), 504–506.

The linear canonical transformations as derived by A. Wintner (*Ann. di Mat.* (4) 13 (1934), 105–112) may also be described as forming the real subgroup of the “complex” (or “symplectic”) group. The algebraic problems associated with the resulting questions in linear dynamics (cf. §153 bis, §154 bis) have been completely solved by J. Williamson, *Amer. Journ. of Math.* 58 (1936), 141–163; 59 (1937), 599–617; 61 (1939), 897–911; [cf. also 62 (1940), 881–911; unfortunately, it was not possible to incorporate his algebraic results into this book.

Chapter III

§155–§158: Historically, the dynamical interest has centered on the quadratic type (1), (7) not only for physical reasons (cf. G. D. Birkhoff, *Dynamical Systems* (1927), 14–32) but also in view of Riemann's n -dimensional differential geometry (cf. §178–§179). Originally, only the reversible case used to be considered. However, it was observed by Levi-Civita (*Torino Atti* 31 (1895), 816–823) that if L is of the form $T + U$ but contains t explicitly, then L can be replaced by a conservative function of the irreversible type (1), provided that t is introduced as an $(n + 1)$ -st coordinate (cf. §9 bis, §93); the corresponding momentum is ignorable in the sense of §182–§183 (the classical transition from (1_1) – (1_3) , §441 to (3_1) – (3_3) , §442 may be thought of as an instance of this procedure). Cf. also E. Cartan, *Leçons sur les invariants intégraux* (1922), *passim*, and G. D. Birkhoff, *op. cit.*, 89–96.

§159–§162: Concerning (14), cf. Jacobi (1845), *Werke* 4, 478–488; A. Wintner, *Quart. Journ. Math. (Oxford)* 7 (1936), 214–218. The oldest instance of the passage from (15_2) to (15_3) is implied by sections 2 and 9 in Book I of Newton's *Principia*. The integral (16) is a slight generalization of one given by Jacobi, who observed that the relation (19_1) , found by Lagrange for $\beta = -2$, holds for any β (cf. the references to §321). The fact mentioned in the second footnote to §159 was pointed out by G. Herglotz, *Seeliger-Festschrift* (1924), 197–199; cf. P. Bohl, *Ztschr. für Math.* 35 (1890), 188–191. Bohl, and then Herglotz, obtained the explicit result of §160 bis by direct integration, instead of using the arbitrariness of the gauge factor. As to §160 and §161, cf. A. Wintner, *Amer. Journ. of Math.* 60 (1938), 473–476.

§163–§164: Cf. L. P. Eisenhart, *Ann. of Math. (2)* 30 (1929), 591–606.

§165–§170: The manifolds of zero velocity occur in a disguised form in Minding's criterion (1838) for the stability of an equilibrium point (cf. the footnote to §134), and were introduced explicitly by Hill (1878) for his case of the restricted problem of three bodies (cf. the references to §462–§476 and §489–§502). Correspondingly, the general rule of §170 (cf. A. Wintner, *Amer. Journ. of Math.* 60 (1938), 471–472) is standard in the Euclidean case of §238.

§171–§181: As to the history of the principle of least action, cf. the comments of P. E. B. Jourdain in no. 167 (1908) of Ostwald's *Klass.*, where detailed references are given. Originally, only the reversible case of a Riemannian geometry (§178–§179) was considered [cf., for instance, a paper of F. Minding (1864; reprinted in *Math. Annalen* 55 (1902), 119–135), which preceded the corresponding considerations of Beltrami and Lipschitz]. Actually, the transition to the general case of §171 is straightforward (cf. Poincaré, *Méth. Nouv.* 3 (1899) 266, and Birkhoff, *Trans. Amer. Math. Soc.* 18 (1917), 203). The useful formal remark of §180 does not seem to be generally known, although it was used by Levi-Civita in his theory of canonical regularization (cf. the references to §398–§399, §415–§420 bis and §446–§454); cf. also G. Darboux, *Comptes Rendus* 108 (1889), 449–450 and P. Painlevé, *Journ. de Math.* (4) 10 (1894), 35–36. The rule of §181, which is fundamental in the theory of surface transformations (cf. H. Poincaré, *Méth. Nouv.* 2 (1893), 370; T. Levi-Civita, *Ann. di Mat.* (3) 5 (1901) 274–278; G. D. Birkhoff, *Dynamical Systems* (1927), 159–162, 210), and was used, e.g., by H. Bruns (*Acta Math.* 11 (1887), 71–73), was obtained already by Jacobi and might occur also in the writings of Hamilton (which are about to be collected in the second volume of his *Math. Papers*).

§182–§183: The possibility of this “reduction by ignorance” becomes apparent if one replaces both the Lagrangian and the Hamiltonian functions by what is called the function of Routh (cf. T. Levi-Civita–U. Amaldi, *Lezioni di Meccanica Razionale* 2₁ [1927], 373–375). The latter function, which exists also when no ignorance of a coordinate (or a momentum) is possible, leads to a mixture of the Lagrangian and the canonical equations, and reduces to L , H in two extreme cases. Cf. also E. R. van Kampen and A. Wintner, *Trans. Amer. Math. Soc.* 44 (1938), 181–182.

§185–§187: It seems to be hard to decide who was the first to write down the energy relation (1₁), which reduces the problem to a quadrature (it must have been known to Euler, but is possibly of an earlier date). As to the qualitative result of this quadrature, cf., e.g., G. Dillner, *Bordeaux Mém.* (2) 5 (1883), 291–304, and P. Stäckel, *Diss.* (Berlin, 1885), 13–17. In order to emphasize the methodical difference between problems in the small and in the large, the discussion in §186 and §187 is purposely based not on this quadrature but on the set of zero velocity.

§188: This procedure of uniformization and expansion is usually attributed to Weierstrass (1866; *Werke* 2, 1–18), although it is contained not only in a posthumous note of Abel (*Œuvres*, 2nd ed., 2, 40–42) but also in Minding's *Handbuch der Theoretischen Mechanik* (1838). The oldest instance of this procedure is the introduction of the eccentric anomaly into the treatment of the elliptic motion (cf. the references to §259).

§189: The linear term following the constant term which is the inverse square root on the right of the approximate formula (11) was considered by P. Fatou, *Acta Astr.* (a) 2 (1931), 135–139; his calculations contain, however, numerical errors. Independently and in a more general direction, a refinement of (11) was recently obtained by Levi-Civita (*Revista Univ. San Marcos (Lima)* 1937, no. 421). An instance of (11) is Newton's result (*Principia*, Book I, Prop. XLV) on the secular precession of the perihelion in case of a non-Newtonian static field of gravitation; cf. also the references to §219.

§194–§198: Originally, Liouville (*Journ. de Math.* (1) 14 (1849), 257–299) arrived at the delimitation of his class of problems by using the method of separation of variables (Jacobi); cf., e.g., §248. The equivalent approach of §194 is more straightforward and is only a particular case of the (isoenergetic) linear U -transformation of Darboux-Painlevé (cf. the references to §180). The determination of more general systems admitting separation of variables actually is a local question in Riemannian geometry; so that the results of the extensive literature of the generalizations of Liouville systems did not seem to belong in this book. It should be mentioned only that the separation of the variables in itself does not solve the dynamical problem, and that the remaining question concerning the “uniformization” of the resulting Abelian inversion problem (cf. §196) is quite unsatisfactory in the usual presentations of the subject. Hadamard (*Bull. des Sci. Math.* (2) 35 (1911), 106–113) has, however, shown how the objections in question can be removed by direct topological discussions. The reduction of this non-local Abelian inversion problem for Liouville systems to the theory of the almost periodic functions, as presented in the text, was given by Wintner (*Amer. Journ. of Math.* 60 (1938), 463–472). The theorem mentioned at the beginning of §198 (H. Bohr, *Medd. Danske Akad.* 10 (1931), no. 12_{IV}; cf. also H. Bohr and B. Jessen, *Pisa Ann.* (2) 1 (1932), 387–398) is analogous to the theorem of §484.

§200–§202: The methodical content of these remarks is to-day commonplace, either because of the whole development of the mathematical literature during the last sixty years, or in view of what may be described as oral tradition.

A fundamental problem, formulated by P. Ehrenfest (*Ztschr. für Physik* 19 (1923), 242–245), is unsolved; cf., in fact, A. Wintner, *loc. cit.*, p. 471.

§203: All this is due to Euler (about 1765); his several papers on the subject and the subsequent literature until 1862 are discussed in the report of Cayley (*Papers* 4, 524–532; references until 1905 are given in Stäckel's article, *Enc. d. math. Wiss.* 41, 497–498).

§205: Cf. J. Andrade, *Journ. de l'Éc. Polytech.* 60 (1890), 55.

§206–§210: The purpose of these and the following articles is to collect in a systematic form certain elementary facts which, even when they are not available in the literature, may nevertheless be considered as known. For an elegant result which depends on Lie's theory, cf. Levi-Civita, *Rend. Acc. Lincei* (5) 5₂ (1896), 164–171. The considerations of §207 can apparently be refined so as to imply that, if $j > k$, the integrals of angular momentum effect a reduction of n to k in case of the problem of k bodies in j dimensions ($k = 2$ in §207; for $k = 3$, cf. W. Ebert, *Astr. Nachr.* 157 (1902), 229–256).

§211–§212: While (12_3) is in Newton's *Principia* (Book I, sections 2 and 3), the energy integral (12_2) seems to be of a later date (cf. the references to §241 and §185). On the other hand, a differential equation of the second order for r alone (cf. §214) was known to Newton. In fact, there results such an equation if one compares (12_3) with Prop. VI, Book I of the *Principia*. This differential equation of the second order (for $1/r$), in which the independent variable is the polar angle, appears explicitly in Clairaut's *Théorie de la Lune* (1765).

§213: This remark was made by Borel (*Nouv. Ann. de Math.* (3) 15 (1896), 236–238). Cf. also the exclusion of the circular paths in §221. Incidentally, Jacobi's calculation of the last multiplier (1845; *Werke* 4, 460) also breaks down in the circular case.

§214: References to the extensive literature dealing with the explicit discussion of paths in case of particular force functions U are given in the reports of Cayley (pp. 516–521) and Stäckel (pp. 494–

496) mentioned before (§203). The dynamical meaning of the second term on the right of (16₂) is explained by section 9 of Book I in Newton's *Principia*.

§215–§219 bis: The question stated in §217 and generalized in §219 was formulated and answered by J. Bertrand, *Comptes Rendus* 77 (1873), 849–853 (as to the subsequent extensive literature, cf. P. Stäckel, *Enc. d. math. Wiss.* 4₁ (1905), 498–499 and P. Liebmann, *ibid.* 3₃ (1914), 526–528). The standard presentation of the subject is such as to need the determination of the second approximation of §219 even for the reduced problem of §217. Actually, the direct considerations of §218 show that the problem of §217 depends only on the first approximation, i.e., on the Jacobi equations, and so it does not involve the lengthy calculations mentioned in §219. It is hard to say why this point is usually overlooked. One reason might be that the topological nature of the problem (cf. §215), or, equivalently, the connection of the problem with existence of an additional integral in the large (cf. §218 bis), is usually not realized; while it is precisely this additional integral which restricts (cf. §148–§149, §151) the characteristic exponents of the Jacobi equations. (The additional integral exists in the case of §219 bis also, but in this case the period and the characteristic exponents are independent of the integration constants; cf. §153.) Another reason seems to be that the trivial characterizations of circular paths, as given in §216, become neglected if one disguises the essential restriction implied by the fact that the problem does not concern arbitrary closed paths but only paths near to a circular solution. It should be emphasized that without this restriction the problem would become extremely difficult, inasmuch as the coefficients of the Jacobi equations are then unknown periodic functions of t , instead of being constants. Incidentally, the proof given in §218 is only a suitable combination of the circular conditions of §216 with Newton's precession formula referred to above (§189). In fact, (21), §218 reduces to Newton's evaluation of the secular precession of the perihelion in case U is a power of r .

§220–§226: These considerations differ only in detail (and caution; cf. (31₁)–(31₂) and §221) from the integration method applied by Jacobi (24th Vorl. ü. Dyn.), and to some extent already by Hamilton (*Phil. Trans.* 1834, 280–281; 1835, 135–139), in case of a static field of radial symmetry.

§228–§232 bis: Cf. G. D. Birkhoff, *Palermo Rend.* 39 (1915), 270–275 or *Trans. Amer. Math. Soc.* 18 (1917), 202–216. In the reversible case, some of the considerations are, of course, of an earlier date. Cf. also D. J. Korteweg, *Sitzber. Akad. Wien* 93 (1886), 995–1040; Lord Kelvin (1891–92), *Papers* 4, 513–522; Sir G. H. Darwin (1897), *Papers* 4, 12–15; also N. Moisseiev, *Rend. Acc. Lincei* (6) 20₂ (1934), 178–182, 256–261, 261–265, 321–327.

§233–§233 bis: Cf. Lord Kelvin, *loc. cit.*; E. T. Whittaker, *M. N. Royal Astr. Soc.* 62 (1902), 186–193, 346–352. Whittaker did not consider the problem of existence; cf. Birkhoff, *loc. cit.* (1917), *passim*, where reference is made to papers of A. Signorini (1912) and L. Tonelli (1911). A systematic account of the relevant investigations of Morse may be found in his *Calculus of Variations in the Large* (1934). For an analysis of certain systems by means of characteristics more elaborate than what is implied by index relations alone, cf. Birkhoff, *Pisa Ann.* (2) 5 (1936), 31–34 and *Mem. Pont. Acad. Novi Lynce.* (3) 1 (1936), Chap. V. In connection with the end of §233, cf. an attempt of L. Vietoris, (*Math. Ztschr.* 19 (1924), 130–135) concerning the “foci” of periodic solutions of the restricted problem of three bodies (these solutions are not algebraic functions of t).

§234–§235: G. W. Hill (1877), *Works* 1, 244–246; H. Poincaré, *Méth. Nouv.* 3 (1899), 280–282. Cf. also the references to §228–§232 bis.

§236–§237 bis: Cf. Wintner, *Sächs. Sitzber.* 82 (1930), 345–354, where it is shown that the Jacobi equation, as given by Poincaré (*Méth. Nouv.* 3, 282–283) for both the reversible and the irreversible cases, is incorrect in the latter case. Another approach, based on a simple conformal mapping, was given by Birkhoff (*loc. cit.*). Cf. also Sir G. H. Darwin, *loc. cit.*, 27–34. As to §237 bis, cf. A. Wintner, *Amer. Journ. of Math.* 53 (1931), 621–622.

§238–§240: E. Strömberg, *Astr. Nachr.* 174 (1906), 33–46; cf. Sir G. H. Darwin, *loc. cit.*, 25–27.

Chapter IV

§241: It would be reasonable to assume that Newton proved (3₂) to be a consequence of (2₁). However, the relevant passages of the *Principia* concern the derivation of (2₁) from Kepler’s laws for circu-

lar planetary motion (and concern, therefore, differentiations instead of integrations). Actually, John Bernoulli (1710; *Opera* 1, 470) appears to have been the first to prove that all paths are conics, if $U(r) = 1/r$. His procedure is, in the main, that described in §214 and is, therefore, identical with the treatment which may to-day be found in elementary text-books; cf. also the references (Newton, Clairaut) to §259. The method of §241, which is so much simpler and is due to Laplace (1798; *Œuvres* 1, 183), seems to be quite forgotten, although it was discovered by Jacobi also (1842; *Werke* 4, 282).

§244: It is a coincidence, which has no historical context, that the names of the three types of conics turn out to correspond to surfaces which have at each of their regular points a second fundamental form of the respective signature (indicatrix of Dupin). In fact, the differential geometry of these surfaces is not even mentioned in the literature.

§245: The geometrical meaning of W was pointed out by P. G. Tait (1865; *Papers* 1, 68–70). As to his paper mentioned there in footnote, cf. *Quart. Journ. of Math.* 7 (1866), 45 (where reference is made to a formula of Hamilton).

§247–§248: In case of parabolic orbits (cf. (15₁), §249), the theorem considered in these articles was first derived by Euler (*Misc. Berol.* 7 (1743), 16). The general theorem was discovered by Lambert (1761) in his monograph, *Insigniores orbitae cometarum proprietates* (no. 133 (1902) of Ostwald's *Klass.*). Lambert's proof consists of lengthy geometrical syntheses. The approach subsequently found by Lagrange (1778; *Œuvres* 3, 559–582) is analytical but still not short. The proof in §248 was given by Jacobi (1837; *Werke* 4, 122); a similar, although longer, proof occurs in the papers of Hamilton (cf. *Phil. Trans.* 1834, 280–286).

§249–§257: The two-fold alternative of §249 for the elliptic case was pointed out by Cayley (1869; *Papers* 7, 387–389). It seems to be difficult to give references to the literature concerning all the constructions described in §250–§257. Actually, the remarks of Jacobi (*Werke*, 1837; 4, 47–48) on conjugate points imply all these constructions, except for the construction of the discontinuous solutions, which were introduced by I. Todhunter, *Researches in the Calculus of Variations* (1871), Chap. VIII. Needless to say, the precise the-

ory of the minimizing orbits considered depends on later developments in calculus of variations; cf., e.g., Ph. Frank, *Monatshefte für Math.* 20 (1909), 171–185 and 189–192.

§259: This elegant method of integration seems to be due to K. Bohlin, *Bull. Astr.* 28 (1911), 144 (certain of its variants are, of course, of a much earlier date; cf., in fact, §261, §267). Correspondingly, the relations obtained by Newton and more explicitly by Clairaut (cf. the references to §211–§212) imply that the function $1/r$ of t is determined, in the case $U = 1/r$, by a linear differential equation of the second order with constant coefficients.

§261–§265: More or less explicitly, all these relations are contained in Book I, Section 3 of Newton's *Principia*, where, of course, the treatment of the three cases of h is rather synthetic and is not always given for all cases.

§268–§269: All this goes back to C. Burrau, *Astr. Nachr.* 135 (1894), 164.

§271: The paths belonging to $U = 1/r^2$ were considered by Newton in his *Principia*, and subsequently discussed in more detail by Cotes; cf. Cayley's report (1862; *Papers* 4, 517).

If the attraction is inversely proportional to an arbitrary, instead of the second, power of the distance and no analytic regularization is possible, it would be desirable to investigate the topological structure of the family of the solution paths near $(x, y) = (0, 0)$. Such a discussion would introduce topological invariants (which must depend very sensitively on the exponent of the force of attraction).

For further references to the literature of the problem of two bodies, cf. G. Herglotz, *Enc. d. math. Wiss.* 6₁, 381–390 (1907), and, as far as the expansions (§274–§299) are concerned, H. Burkhardt, *ibid.* 2₁, 827–829, 891–902, 1345–1349 (1912) and W. F. Osgood, *ibid.* 2₂, 44–47 (1901).

§278: Lagrange (1771), *Oeuvres* 3, 113–138; Bessel (1824), *Ges. Abh.* 1, 84–102.

§279–§280: Bessel (1818; 1824), *ibid.*, 1, 17–20; 100.

§281–§282: Cf. Burkhardt, *loc. cit.*, pp. 825–827, 892–895.

§283–§284: The first correct approach to (44₁), i.e., to (44 bis), is due to Carlini (1817; cf. Jacobi (1850), *Werke* 6, 188–245) whose

work remained, however, unnoticed until Jacobi (1848; Werke 6, 175–188) freed it from errors of calculation. Laplace (Œuvres 5, 473–489) arrived at (44₁) by using considerations which were published (1827) after his death and which he realized (*ibid.*, p. 489) to be heuristic; in fact, he proved the asymptotic formula for imaginary, but would need it for real, values of the argument (in this connection, cf. A. Wintner, *Proc. Nat. Acad. Wash.* 20 (1934), 57–62; P. Hartman, *Amer. Journ. of Math.* 62 (1940), 115–121). The relation (44₂) is more recondite than (44₁) and was not considered by Laplace but only by Carlini; cf. Jacobi, *loc. cit.* According to Cauchy (1843; Œuvres (1) 12, 164), who derived (45₁), both (44₁) and (44₂) may be obtained simply by his complex function-theoretical method (1843; Œuvres (1) 8, 128–133 and 1845; Œuvres (1) 9, 75–83); this fact was rediscovered and simplified by Riemann (1863 (1876); Werke, 2nd ed., 426–430). For a modernized presentation of this “method of steepest descent,” cf. O. Perron, *Münch. Sitzber.* 1917, 191–220, where (45₂) is proved also. The introduction of the number $0.6 \dots$, defined by (48), is due to Laplace (*loc. cit.*); as to its value (49), cf. a letter (1889) of Stieltjes to Hermite (*Correspondence*, 1, 433–434).

Further references to §277–§284 may be found in Watson’s *Treatise on Bessel Functions* (1922) and in Burkhardt’s report, *Jahresber. d. D. M. V.* 10₁ (1908), Chap. III. The importance of the problems of §283–§299 in the historical development of the theory of analytic functions is discussed in the report of Brill and Noether, *ibid.* 3 (1894), Chap. II.

§285–§299: Lagrange introduced his solution rule in 1770 (Œuvres 3, 126) and then (1771) applied it to Kepler’s equation (*ibid.*, 113–138). In view of his formal rearrangements of series, the treatment presented in §287–§288 may be thought of as a modernization of his approach (cf. §297–§298). The standard proof of (53₁)–(53₂) is not this but the one described in §291–§292, (cf. e.g., Tchebycheff’s Œuvres 1, 251–270 [1857], or Puiseux’s note in Lagrange’s Œuvres 12, 341–346), as discovered by Cauchy (1829; Œuvres (1) 2, 41–48) in his theory of analytic functions (for further references in this direction, cf. Brill-Noether, *loc. cit.*, 176–179, 187–189 and Osgood, *loc. cit.* 46–47). The critical remark at the end of §292 is, of course, of a later date (1906; A. Hurwitz, Werke 1, 655–659). The results of §294–§295 were found by C. L. V. Charlier, *Lund Obs. Medd.* no. 22, and by Levi-Civita, *Rend. Acc. Lincei* (5) 13₁ (1904), 260–268; actually, the inequality (68) was discovered by Kapteyn (*Ann. Éc.*

Norm. Sup. (3) 10 (1893), 96–99), who also recognized its rôle for Kepler's equation (cf. also Watson, *op. cit.*, 268–270 and Chap. XVII). As to the analogue of the expansions of §295 in case of hyperbolic motions, cf. H. Block, *Ark. för Mat. Astr. Fys.* 1 (1904), 467–479.

The importance of the considerations of §300–§312 bis lies in their rôle of supplying the elementary approximation to the restricted problem of three bodies. For instance, (7₂), §300 explains the observation made by Jacobi after his formula (11) of his 5th Vorl. ü. Dyn. Correspondingly, the explicit rules of §302–§303 may be useful in connection with the ring transformation considered by H. Poincaré (*Acta Math.* 13 (1890), 171–174; *Méth. Nouv.* 3 (1899), 196–200, 374–381; *Palermo Rend.* 33 (1912), 375–407) and by G. D. Birkhoff (*Palermo Rend.* 39 (1915), 288–295; cf. *Trans. Amer. Math. Soc.* 14 (1913), 14–22, *Acta Math.* 47 (1926), 297–311; *Dynamical Systems* (1927), Chap. VI). As to the arrangements of §307–§309 and §312–§312 bis, cf. A. Wintner, *Math. Ztschr.* 34 (1932), 367–373.

§305–§307: The conditions discussed are needed in the theory of the periodic solutions of the restricted problem of three bodies. Cf. the more advanced parts of Poincaré's works just mentioned, and his papers in *Bull. Astr.* 1 (1884), 65–74, 8 (1891), 12–24, 19 (1902), 177–198; T. Levi-Civita, *Ann. di Mat.* (3) 5 (1901), 284–289; G. D. Birkhoff, *Palermo Rend.* 39 (1915), 295–313, *Pisa Ann.* (2) 4 (1935), 267–306, and B. O. Koopman, *Trans. Amer. Math. Soc.* 29 (1927), 310–331; P. Stäckel, *Jahresber. d. D. M. V.* 28 (1919), 180–181; A. Wintner, *Sächs. Sitzber.* 82 (1930), 3–56; *Math. Annalen* 96 (1926), 284–318, and M. Martin, *Amer. Journ. of Math.* 53 (1931), 259–273; E. Hölder, *Sächs. Sitzber.* 83 (1931), 179–184, *Amer. Journ. of Math.* 60 (1938), 801–814 and *Math. Ztschr.* 31 (1929), 225–239 (cf. L. Lichtenstein, *ibid.* 17 (1923), 62–110); also T. Uno, *Sendai Astr. Rap.* 1 (1938), 149–191.

§310–§311: T. Levi-Civita, *Ann. di Mat.* (3) 9 (1904), 21–25; cf. also F. R. Moulton, *Proc. London Math. Soc.* (2) 11 (1913), 367–384, where reference is made to the work of C. Burrau (cf. §268–§269 above).

Chapter V

References to the classical literature of the problem of several bodies may be found in the following text-books: O. Dziobek, *Die mathematischen Theorien der Planeten-Bewegung*; 1888 (the page num-

bers given below refer to the American edition, 1892); F. Tisserand, *Traité de Mécanique Céleste*, 1, 1896; H. C. Plummer, *An Introductory Treatise on Dynamical Astronomy*, 1918.

A rather useful bibliography is due to R. Marcolongo, *Il problema dei tre corpi da Newton (1686) ai nostri giorni* (no. 403–405 (1919) of the *Manuali Hoepli*).

§313: This formal approach to the “physical” problem is not, of course, that of Newton, and is formulated along lines influenced by Mach’s critique. The astronomical issues involved are discussed by E. Arndt, *Enc. d. math. Wiss.* 6₁, 3–15 (1905) and J. Bauschinger, *ibid.* 843–895 (1919). The discovery of the force function $\{ \}$ in (1) is due to Lagrange (1773; *Œuvres* 6, 348, also 1777; 4, 408).

§315–§320: Although the actual content of the ten classical integrals was known not later than the end of the first half of the 18th century (cf. the comments of P. E. B. Jourdain on Newton, Clairaut, d’Arcy, D. Bernoulli and Euler in no. 191 (1914) of Ostwald’s *Klass.*), their present form and the discovery of the formulation (7₁) of (7₂) are due to Lagrange (cf., e.g., *Œuvres* 9, 386 and *Œuvres* 6, 240, where (7₁) is given for $n = 3$). The fundamental observation that the integrals of §316–§317 are necessitated by the Galilei automorphisms of the equations of motion appears in Jacobi’s *Vorl. ü. Dyn.* (1842) but must have been known to Lagrange also (1777; *Œuvres* 4, 406), at least implicitly. The embedding of §315–§317 into the general theory of Lie is discussed, e.g., by F. Engel, *Gött. Nachr.* 1916, 270–275, 1917, 189–198. As to §319, cf. also J. R. Schütz, *Gött. Nachr.* 1897, 110–123. The completeness of the Galilei group, as proved in §318, is usually considered as evident (which it is not; cf. A. Wintner, *Amer. Journ. of Math.* 60 (1938), 473–476). The arbitrariness of the gauge factor of §315 bis (cf. Jacobi (1845), *Werke* 4, 485–488; explicitly formulated by O. Dzio-bek, *op. cit.*, p. 64) is only an instance of the Galilei-Newton principle of dynamical similarity.

§320 bis: H. Bruns, *Sächs. Sitzber.* 13 (1887), 1–39, 55–82 (= *Acta Math.* 11 (1887), 25–96); H. Poincaré, *Méth. Nouv.* 1 (1892), 233–334; P. Painlevé, *Comptes Rendus* 124 (1897), 173–176, *Bull. Astr.* 15 (1898), 81–113, *Comptes Rendus* 130 (1900), 1699–1701. A slip in the work of Bruns was corrected by Poincaré, *Comptes Rendus* 123 (1896), 1224–1228. The attitude taken in §320 bis with regard

to algebraic integrals may be quite unorthodox but is certainly necessitated by any geometrical, i.e., non-local, concept of a non-integrable dynamical system.

In this connection, cf. T. Levi-Civita, *Verh. des III. Int. Math. Kongr. 1904* (1905), 407–408 and his report in *Comptes Rendus du 2me Congr. Int. de Méc. Appliquée*, 1926 (1927); cf. also J. Chazy *Bull. Astr.* (2) 8 (1933), 403–436.

§321: Cf. Wintner, *loc. cit.* The integrals (17) were pointed out by Jacobi (4th Vorl. ü. Dyn.), who has also shown that (17) reduces the rectilinear motion of $n = 3$ bodies to quadratures (1837, 1844; *Werke* 4, 481–488, 533–539).

§322 bis: Lagrange (1772), *Œuvres* 6, 233–240 (where $n = 3$).

§323: Laplace (1798), *Œuvres* 1, 65–69 (cf. 3, 173), where $C = 0$ is excluded.

§324–§331 bis: In the literature, these kinematical facts are not stated and proved in a systematic form, although most of them cannot be considered as “evident” (cf. §373 bis, §374 bis). The formulae of §325 bis suggested the notion of a flat solution, as introduced in §325. While this notion is superfluous if $n = 3$, it shows its usefulness if one attempts to generalize for an arbitrary n certain results which are classical for $n = 3$. This is illustrated by the result of §326, which in the literature occurs only in the somewhat misleading case $n = 3$ (treated first by O. Dziobek, *op. cit.*, p. 63, and then more simply by K. Sundman at the beginning of his paper referred to below). Other instances are supplied by the theory of homographic solutions (§373–§374), where the main theorems implicitly depend on the notion of a flat solution (although the content of these theorems may be found in the literature). The result of §327 is astronomical tradition, at least if $n = 3$. While a similar remark must apply to §328–§329, the result of §331 is due to P. Pizzetti (*Rend. Acc. Lincei* (5) 13₁ (1904), 24–25, where n is arbitrary; it seems to be hard to locate an earlier reference even for $n = 3$). The straightforward construction described in the footnote to §325 is due to a recent conversation with Dr. E. R. van Kampen.

§332–§332 bis: These fundamental consequences of Lagrange’s identity (2₄) were drawn by Jacobi, 4th Vorl. ü. Dyn. (1842).

A mistaken explanation of a paradox of Jacobi (*loc. cit.*) on colli-

sions was given by H. Seeliger (*Astr. Nachr.* 113 (1885), 358), a correct one by E. Freundlich (*ibid.*, 208 (1919), 209–212).

Essential refinements of the considerations of §332–§332 bis are due to the investigations of J. Chazy, which were announced in the *Comptes Rendus* and then collected in his paper, *Ann. Éc. Norm. Sup.* (2) 39 (1922), 29–130. [The corresponding questions in the limiting case of the restricted problem of three bodies were subsequently considered by B. O. Koopman, *Trans. Amer. Math. Soc.* 29 (1927), 288–304.] Chazy first proves that if h is positive, the ratio of the least and of the greatest of the $\frac{1}{2}n(n-1)$ mutual distances tends to a limit as t tends to infinity, and that this limit is a continuous function of the initial conditions. Chazy then classifies, for $n = 3$, the different solutions of positive h in terms of the order of magnitude (for large t) of the mutual distances. He arrives at corresponding results in the limiting case $h = 0$ also. Finally, he develops the beginnings of a corresponding classification theory also in case of a negative energy (which is the most difficult case; cf. the parenthetical remark of §332 bis). In his paper *Journ. de Math.* (3) 8 (1929), 353–380, and in his report *Bull. Astr.* (2) 8 (1933), 403–436 on his theory of classification, Chazy succeeded in obtaining further results in this direction. Unfortunately, the proofs of his deep results turned out to be too lengthy for a detailed presentation in this book. Cf. also the references to §431–§431 bis.

§333–§338 bis: Although the presentation is slightly simplified in the text, all these results and methods are due to K. F. Sundman, *Acta Soc. Sci. Fenn.* 35 (1909), no. 9, where $n = 3$; his considerations hold, however, for any n , as has been observed by H. Block (*Lund Astr. Obs. Medd.* (2) 6 (1909), no. 6) and rediscovered by J. Chazy (*Bull. Astr.* 35 (1918), 321–341; cf. *Comptes Rendus* 157 (1913), 688–691). It turned out after the publication of Sundman's paper, that his preliminary result, $C = 0$, (§335) was known to Weierstrass (letter (1889) to G. Mittag-Leffler; *Acta Math.*, 35 (1912), 57–58). This fundamental paper of Sundman has attracted essentially less attention than his theory of binary collisions (it was not even reviewed in the *Fortschr. d. Math.*, and subsequently it was not reproduced in *Acta Math.* 36 (1913); cf. §348–§352 below).

It appeared to be convenient to defer the formulation of the actual content of these results until §361–§364.

The distinctly Tauberian character of Sundman's considerations,

which imply the corresponding $(C, 1)$ -results of somewhat later date (Hardy-Littlewood), was only recently pointed out by A. Wintner (cf. R. P. Boas, Jr., *Amer. Journ. of Math.* 61 (1939), 161–174; subsequently, J. Karamata (*ibid.*, 769–770) has shown that Sundman's Tauberian condition concerning unilateral boundedness may be replaced in the usual manner by the corresponding condition on oscillation). It is interesting that also one of the oldest Tauberian theorems, namely, that of §362, was introduced by Hadamard in connection with a dynamical question (*Journ. de Math.* (5) 3, 334; for a refinement in terms of an absolute constant, cf., e.g., E. Landau, *Proc. London Math. Soc.* (2) 13 (1914), 43–49).

§339: Cf. J. Chazy, *Ann. Éc. Norm. Sup.* (3) 39 (1922), 124. For $n = 3$, Chazy (*ibid.*, 124–126; *Comptes Rendus* 157 (1913), 1398–1400) proves a corresponding, though weaker, theorem for binary collisions, by showing that the distance between two of the bodies cannot tend to zero when t tends to infinity, if at the same time their distances from the third body exceed a positive lower bound.

§340–§343: The heliocentric equations (12) are as old as the beginnings of the theory of perturbations and must, therefore, have been standard by the end of the first half of the 18th century. The introduction of the disturbing function, (11₂), is of a later date, since it was made possible only by Lagrange's introduction of (3₂), §314.

§344–§347 bis: These particular solutions are due to A. E. Fransen, *Öfv. Stockh. Akad.* 52 (1895), 783–805. Cf. also J. Chazy, *Comptes Rendus* 169 (1919), 526–529, *Bull. Astr.* (2) 1 (1921), 171–188.

§348–§352: This theory is due, in its present form, to Sundman (*Acta Soc. Sci. Fenn.* 34 (1907), no. 6; reproduced in *Acta Math.* 36 (1912), 105–179), although several results were known before him (Bruns, Painlevé; also Weierstrass); cf. the references to §407–§412. An attempt of G. Bisconini (*Acta Math.* 30 (1904), 49–91) failed, inasmuch as he had to postulate a result which is equivalent to that proved in §352. Although Sundman considered only the case $n = 3$, the transition to any n is straightforward, at least if his treatment of $n = 3$ is simplified, as above, at some unessential points.

§353–§354: These facts agree with the astronomical tradition but were first proved by J. Chazy (*Comptes Rendus* 168 (1919), 81–83; cf. *Ann. Éc. Norm. Sup.* (3) 39 (1922), 127).

§355–§360: The central configurations belonging to $n = 3$ were discovered in the collinear case (§358) by Euler (*Nova Comm. Petrop.* 11 (1767), 144–151; *Hist. de l'Acad. Berl.* 1770, 194–220), in the case of §359 by Lagrange (1772; *Œuvres* 6, 272–292), who also arrived at Euler's case. Incidentally, Euler derived his quintic equation (by a direct consideration) for the limiting case of the restricted problem also. The general approach to central configurations (§355, §357), as applied in §358–§359, is due to O. Dziobek (*Astr. Nachr.* 152 (1900), 33–46). Actually, the notion of a central configuration was introduced by Laplace (1789; *Œuvres* 11, 553–558 = 1805; 4, 307–313), who was led to it by his straightforward, but rather incomplete, treatment of Lagrange's homothetic solutions (cf. below). It is curious that most of the elementary text-books, and even Cayley's otherwise very useful historical report (1862; *Papers* 4, 540), attribute these solutions to Laplace (who, for his part, did not have the habit of giving references; in regard to this chapter in the history of celestial mechanics, E. T. Bell's *Men of Mathematics* is not much overdone). Dziobek's fundamental paper is not usually mentioned in the literature [cf., e.g., H. Andoyer, *Bull. Astr.* 23 (1906), 50–59; F. R. Moulton, *Ann. of Math.* (2) 12 (1910), 1–17; W. D. MacMillan and W. Bartky, *Trans. Amer. Math. Soc.* 34 (1932), 838–875; also W. L. Williams, *ibid.*, 44 (1938), 562–579, where the non-collinear planar case of $n = 5$ bodies is considered]. In particular, Dziobek arrived at (13) and at several further results for the case $n = 4$, and formulated a conjecture subsequently discussed in detail by MacMillan and Bartky (*loc. cit.*). Dziobek's paper was preceded by a note of R. Lehmann-Filhès (*Astr. Nachr.* 127 (1891), 137–144), who observed the configuration of §359 for $n = 4$ and considered the case (i) of §360 for any n (as to the latter case, cf. also F. R. Moulton, *loc. cit.*, where the discussions, based on the approach described in §356, depend on a determinant treated by T. H. Hilbrandt). The calculations connected with the known configurations mentioned under (iii), §360 are, of course, of a trivial nature; cf., e.g., R. Hoppe, *Arch. der Math.* 64 (1879), 218–223, Emilia Breglia, *Giorn. di Mat.* (3) 7 (1916), 165–168.

§361–§364: Cf. the references to §333–§338 bis.

§365–§368 bis: Difficulties of this type (cf. also §411, §425) were first recognized by P. Painlevé; cf. his *Leçons sur la théorie analytique des équations différentielles* (Stockholm, 1895), Paris,

1897, pp. 543–577 and 587–589 (where reference is made to a consideration of Poincaré). A note of H. von Zeipel (*Ark. för Mat. Astr. Fys.* 4 (1908), no. 32) indicates a consideration to the effect that, if U becomes infinite when t tends to a finite value, then J must tend to infinity, unless all bodies tend to definite limiting positions. But it seems to be hard to fill in the gaps. A consideration of E. Freundlich (*Berl. Sitzber.* 1918, 168–188) seems to overlook the actual difficulties. Their appearance in the problem of simultaneous collisions was further discussed by J. Chazy, *Bull. Astr.* 35 (1918), 321–389. According to Chazy (cf. loc. cit. 341–364), the contingency of §368 concerning spirals is certainly impossible if $n = 3$. Actually, no case is known in which this contingency occurs.

§369–§378: The result of §373 is due to Pizzetti (*Rend. Acc. Lincei* (5) 13₁ (1904), 276–283), that of §374 for any n to Pizzetti (*ibid.*), for $n = 3$ to Lagrange (1772; *Œuvres* 6, 272–292, where it is emphasized (p. 292) that this is the central theorem; the approach of Laplace, mentioned above, disregards this theorem completely). In particular, §377 goes back to Lagrange's work on $n = 3$; cf. Dziobek, loc. cit. The complete verifications of §375–§377, which are usually omitted in the literature, had to be included, since otherwise it is hardly possible to prove that all cases enumerated in §378 actually exist. The example of §374 bis was given by T. Banachiewicz, *Comptes Rendus*, 142 (1906), 510–512, his considerations are made somewhat difficult, however, by the fact that he neither mentions nor uses the isosceles character of these solutions (cf. A. Wintner, *Amer. Journ. of Math.* 60 (1938), 473); this might be the reason that the example of §373 bis, for which a pure calculation without any geometrical limitations would be still more involved, does not occur in the literature.

§379–§380: These are the solutions “stationary in sense of Routh” (T. Levi-Civita, *Prace Mat.-Fyz.* 17 (1906), 1–40). Cf. also H. Andoyer, *Bull. Astr.* 23 (1906), 50–59.

§381–§382: As to §381, cf., e.g., H. Andoyer, *ibid.*, 129–146. The results (I) and (II) of §382 are due to J. Liouville (*Journ. de Math.* (1) 7 (1842), 110–113; (2) 1 (1856), 248–264) and to G. Gascheau (*Thèse*, Paris, Bachelier, 1843; *Comptes Rendus* 16 (1843), 393–394), respectively. The subsequent results of H. Gylden (*Bull. Astr.* 1

(1884), 361–369) and H. C. Plummer (*M. N. Royal Astr. Soc.* 62 (1902), 6–17) on the limiting case of the restricted problem (cf. §476 below) may be thought of as contained in the results of Gascheau and Liouville, respectively.

§382 bis: J. C. Maxwell (1856), *Papers* 1, 288–376 (Part II); cf. L. Lichtenstein, *Math. Ztschr.* 17 (1923), 62–110 and *Pisa Ann.* (2) 1 (1932), 173–213.

§383–§388: As to the introduction of suitable linear combinations of barycentric coordinates in general, cf. P. Pizzetti, *Atti Acc. Torino* 38 (1903), 954–961. The remarks of §384 are due to Poincaré (*Bull. Astr.* 14 (1897), 53–67; reprinted in *Acta Math.* 21 (1897), 83–97). The ideal masses of §385 and the corresponding coordinate chains, together with their elegant consequences (15₁)–(15₃), were introduced for $n = 3$ by Jacobi (1842; *Werke* 4, 299–306) and subsequently extended for any n by R. Radau (*Ann. Éc. Norm. Sup.* (1) 5 (1868), 311–375); cf. also F. Hopfner, *Astr. Nachr.* 195 (1913), 256–262. The geometrical interpretation of the conservation of the angular momentum of $n = 3$ bodies (§388) was also pointed out by Jacobi (*loc. cit.*, 307–308). That there are exceptional cases in which this interpretation fails, does not seem to be mentioned in the literature. The problem itself, as formulated in §388 bis, is not likely to be an easy one. The fundamental fact stated in §389 was proved by W. D. MacMillan (in a paper of E. J. Wilczynski, *Ann. di Mat.* (3) 21 (1913), 17–31); cf. also the presentation of J. Chazy, *Bull. Astr.* (2) 1 (1921), 171–188. In the literature, the corresponding problem for $n > 3$, as formulated in §389 bis, is not considered, since it depends on the notion of a flat solution.

§390–§397: The theory of reduction of the problem of n bodies goes back to Lagrange (1771; *Œuvres* 6, 227–331), who proved that the classical integrals reduce the general problem of $n = 3$ bodies to a system of the seventh order (cf. §434). Lagrange's paper appears to have escaped the attention of Jacobi (1842; *Werke* 4, 295–314), who arrived at the same result with the help of his considerations referred to before (§387–§388). Jacobi's celebrated "elimination of the node," though not in his straightforward geometrical form, is contained in the formulae of Lagrange (however, neither Lagrange nor Jacobi arrived at a canonical form of the reduced equations of motion). The subsequent literature of the subject is quite extensive

and is discussed on pp. 29–44 of Marcolongo's report. The last approach considered there, that of Levi-Civita (*Atti Ist. Veneto* 74 (1915), 907–939), was afterwards presented in another form by Maria Ronchi (*ibid.*, 76 (1917), 1221–1225). Cf. also E. R. van Kampen and A. Wintner, *Amer. Journ. of Math.* 59 (1937), 153–166; 269, where the reduction is symmetric in the $n = 3$ masses. A rather geometrical approach to Lagrange's reduction is due to G. D. Birkhoff (*Dynamical Systems*, 1927, 283–288); his considerations are based directly on the 18-dimensional Cartesian phase space (cf. §390–§392 for $n = 3$), in which he follows the flow consisting of the solution paths which constitute the intersection of the hypersurfaces formed by the ten classical integrals. In this connection, cf. E. Cartan, *Leçons sur les invariants intégraux*, 1922, 172–181, where the problem of reduction is interpreted kinematically, from the point of view of the infinitesimal transformation involved. In the literature, the H of the reduced problem does not occur in the form (33), §394. However, the latter may be obtained by subjecting the H given by van Kampen and Wintner (*loc. cit.*) to the binary substitution which is the canonical extension of the third of their equations (51). The introduction of this substitution seemed to be advisable for reasons which are apparent from §435. Birkhoff's reduced flow then follows (§437–§440) in terms of differential equations which are symmetric in the masses, intrinsic, and more or less explicit. However, this approach to the reduced model depends on a fascinating unsolved problem, formulated in §436 (the singular cases in question are, of course, rather exceptional).

§398–§399: Besides the literature covered in chap. II of Marcolongo's bibliography, cf. the studies of Levi-Civita (*Rend. Acc. Lincei* (5) 24₂ (1915), 61–75, 235–248; 421–433, 485–501, 553–569) on the planar case of $n = 3$, and Cartan's *Leçons* (*loc. cit.*), finally a note of F. D. Murnaghan, *Amer. Journ. of Math.* 58 (1936), 829–832, where a short deduction of (37), §399 is given. It would be interesting to calculate (cf. W. Kaplan, *Compositio Math.* 5 (1938), 327–346), at least in some cases (first of all for $n = 3$, $C = 0$), the principal topological characteristics of the algebraic manifolds representing the intersections of the classical integral surfaces.

A qualitative investigation of the rectilinear case of $n = 3$ (in the reduced form given by Euler (*Nova Acta Petrop.* 3 (1776), 126–141 and then (1845) by Jacobi, *Werke* 4, 478–485)) is due to J. Chazy,

Bull. Soc. Math. de France *55* (1927), 222–268. This case, which does not have much astronomical interest, is to-day the only one in which a detailed qualitative study (cf. G. D. Birkhoff, *Dynamical Systems* (1927), 288–291) can be attempted.

§399 bis–§402: The formal simplifications arising for $h = 0$ are implied by the general remarks of Jacobi (loc. cit., 485–488) and were, in the rectilinear case of $n = 3$ bodies, recognized already by Euler (loc. cit.). The oldest instance of this simplification is the integration of the problem of two bodies in case of parabolic motion, when compared with the more complicated case of elliptic or hyperbolic motion. On p. 65 of his book, Dziobek makes a statement concerning the case in which also $C = 0$. In this connection, cf. Cartan, op. cit., 181–185. As to the approach of §399 bis–§400, cf. A. Wintner, *Quart. Journ. Math. (Oxford)* *7* (1936), 214–218. The remarks of §401–§402 might clear up a note of W. Ebert, *Comptes Rendus* *131* (1900), 251–253.

§403–§406: The observation as to the existence of a centre of force for $n = 3$ must have been known implicitly to Laplace (1789; *Œuvres* *11*, 554–555), but seems first to occur in a note of J. Hargrave, *Phil. Mag.* (4) *16* (1858), 466–473. It would be a mistake to expect that the remarks of §405 reduce the problems of §374–§374 bis and §389 to elementary kinematical discussions. The quintic equation of §406 was calculated only for the sake of completeness; its kinematical meaning, if any, is not known.

§407–§412: All these results are due to Painlevé (pp 569–577, 582–586 of his *Stockholm Leçons*, referred to above (§365–§368 bis), and *Comptes Rendus*, *123* (1896), 636–639, 871–873; cf. also *139* (1904), 1170–1174). Some of his results for $n = 3$ were apparently known to Weierstrass; cf. the letter (1889) referred to above (§333–§338 bis). The introduction of (9), §414 is mentioned already by Bruns (*Astr. Nachr.* *109* (1884), 219–220).

§415–§420 bis: The result of §420 is stated to be true already by Bruns (loc. cit.). It was known to Weierstrass also (loc. cit.). However, the first proof available in the literature is due to Sundman (cf. the references to §348–§352). His calculations are quite involved, apparently because no use is made of the canonical form of the differential equations. The fundamental canonical transformation of §50 and the elegant approach of §415–§419, which does not

sacrifice the dynamical form of the equations, were discovered by Levi-Civita (*Comptes Rendus* 162 (1916), 625–628; cf. *Acta Math.* 42 (1920), 99–144).

§421–§424: All this is due to H. Block (*Ark. för Mat. Astr. Fys.* 5 (1909), no. 9; cf. *Lund Astr. Obs. Medd.* (2) 6 (1909), no. 6). Block's theory was rediscovered by J. Chazy, who considered (*Bull. Astr.* 35 (1918), 341–364), in addition, the question of completeness, mentioned in §421 bis. As to the footnote in §423, cf. H. Poincaré (1879), *Œuvres* 1, pp. XCIX–CXXIX; *Acta Math.* 13 (1890), 27–41.

§425: This extension of the binary case is obvious. As to the remaining cases, cf. the references to §365–§368 bis.

§426–§430: In the literature, the treatment of the question dealt with in §427–§429 is quite indirect, since it is made to depend on the deeper theorem formulated in §431 (which, incidentally, excludes the case $C = 0$ of §431 bis). However, it seemed to be advisable from the methodical point of view, to keep the theorem of §431 in the background, by presenting a direct approach (§427–§430) to the simpler fact formulated at the end of §426. A further possible simplification, now contained in §429, was pointed out by Dr. E. R. van Kampen.

§431–§431 bis: The results of §431 bis for $C = 0$ follow, though quite indirectly, from the investigations of J. Chazy, mentioned at the end of the references to §365–§368 bis; cf., in fact, loc. cit., pp. 382–383. The theorem for $C \neq 0$, mentioned in §431, is due to Sundman (cf. his papers referred to in connection with §348–§352). Actually, his proof contains an error which, however, was easily removed by Hadamard (*Bull. des Sci. Math.* 39₁ (1915), 249–264) along the line of Sundman's ideas. These ideas represent essential refinements of the considerations of Jacobi (cf. §332–§332 bis). In fact, it is now of no avail simply to let t tend to infinity, since explicit estimates of the distances are needed along finite t -intervals which cluster at $t = \infty$. In this sense, the theorem of §431 may be thought of as being of the same Tauberian nature as the result of §337–§338 bis (although the distinctly Tauberian part of the considerations has not hitherto been isolated in the form of a general lemma on real functions). The Sundman-Hadamard technique of the estimations involved was further developed by Chazy, *Ann. Éc. Norm. Sup.* (2) 39 (1922), 109–126, and by Birkhoff, *Dynamical*

Systems, 1927, 275–283 (also 291–292; cf. J. J. L. Hinrichsen, *Trans. Amer. Math. Soc.* 36 (1934), 306–314).

A paradigm of the results available to this method of detailed estimates may be formulated as follows: If the $n = 3$ masses and the integration constants $h < 0$, $C \neq 0$ are fixed, there exists a sufficiently small positive number with the property that, if $J = J(t)$ becomes less than this number for some t , then two of the mutual distances must tend with t to infinity, while the third remains under a fixed upper bound; in addition, it is always the same body which is relatively remote throughout the entire motion. Cf. Birkhoff, *loc. cit.*

§432–§440: The methodical points of view taken in these articles were greatly influenced by repeated discussions with Professor G. D. Birkhoff. The relation of §433–§440 to the literature of the subject may be seen from the references to §390–§397 (cf., in particular Birkhoff, *loc. cit.*). The expansions described in §432–§432 bis were established by Sundman, *loc. cit.* (In this connection, cf. H. Poincaré (1886), *Œuvres* 1, 181–189; P. Painlevé, *Stockholm Leçons*, 577–582; also posthumous (1857) notes of Cauchy, *Œuvres* (1) 12, 445–455; finally, the statements of Bruns and of Weierstrass, referred to in connection with §415–§420 bis). Typical of the usual emphasis laid on the formulation mentioned in §432 bis are the remarks of É. Picard, *Bull. des Sci. Math.* (2) 37₁ (1913), 313–320. On the other hand, the astronomers were from the beginning more than sceptical concerning the usefulness of Sundman's expansions. As to the end of §432 bis, cf. the calculations of D. Belorizky (e.g., *Bull. Astr.* (2) 6 (1930), 417–434).

Chapter VI

§441–§443: Euler's second lunar theory, which is based on the introduction of the rotating coordinate system and on the model of §441, was published in 1772 in a monograph (*"Theoria motuum lunae . . ."*). Jacobi, who rediscovered this model in 1836 (*Werke* 4, 37–38), apparently recognized its relevance for the theory of minor planets also, and pointed out the integral (7₄). For further references, cf., e.g., Newcomb's report on lunar theory, *Atti del IV. Congr. Int. Mat.* 1908, 1; 135–143.

§443 bis: Cf. Sir. G. H. Darwin (1897), *Papers* 4, 4.

§444: Cf. the references to §203.

§444 bis: This remark is contained in the calculations of H. Samter, *Astr. Nachr.* 217 (1922), 129–152, although he does not mention that the model of two fixed centra (Euler) is now actually refined by the inclusion of the centrifugal terms, and that in this respect the case of two equal masses is exceptional.

§445: The proof for the non-existence of new integrals of the type described in the second part of §320 bis first was given by Poincaré for the restricted problem (*Acta Math.* 13 (1890), 259–265; cf. *Méth. Nouv.* 1 (1892), chap. V). Recently, C. L. Siegel (*Trans. Amer. Math. Soc.* 39 (1936), 225–233) transferred to the restricted problem the results of Bruns concerning the non-existence of new algebraic integrals (cf. §320 bis).

§446–§454: The transformation applied in §451 is precisely that by means of which Euler integrated his problem of two fixed centra (cf. §203). After Burrau's fundamental papers (*Astr. Nachr.* 135 (1894), 233–240; 136 (1894), 161–174; cf. also his report *Astr. Ges. Vjs.* 33 (1898), 21–23 on Darwin's calculations), which started out from a numerical question formulated by T. N. Thiele (1892) as a prize problem of the Danish Academy, Thiele has shown (*Astr. Nachr.* 138 (1896), 1–10) that Euler's substitution supplies a regularization of the restricted problem also. Actually, Thiele considered (*loc. cit.*) only the case of equal masses (cf. §452). The extension of his regularization to the case of arbitrary masses (§451) is due to Burrau (*Astr. Ges. Vjs.* 41 (1906), 261–266; cf. Levi-Civita, *Rend. Acc. Lincei* (5) 24₂ (1915), 553–559). However, somewhat before this paper of Burrau, and without knowing Thiele's treatment of the symmetric case, Levi-Civita (*Verh. des III. Int. Math. Kongr.* 1904 (1905), 402–408; *Acta Math.* 30 (1904), 305–327) discovered the simpler (and, though only local, in principle equivalent) regularization given in §447–§451. A simple description of a collision in terms of Levi-Civita's coordinates is mentioned by Birkhoff, *Pisa Ann.* (2) 4 (1935), 272–273. The regularization of §453 was introduced by Birkhoff in order to facilitate topological discussions (*Palermo Rend.* 39 (1915), 276–288). As to §454, cf. Wintner, *Math. Ztschr.* 32 (1930), 691–698.

The majority of the numerical investigations mentioned in §452 concern families of periodic (and asymptotic) solutions, and are due to E. Strömgren and his collaborators. The list of publications of these investigations is quite extensive, and may be found in Ström-

gren's reports. The most complete of these reports is in *Bull. Astr.* (2) 9 (1933), 87–130, where the whole field is reviewed in a comprehensive way. Cf. also the references to §519 bis below.

§455–§461: A detailed treatment of the essential problem considered here seems to be missing in the literature. As pointed out by G. Hamel in the *Fortschr. d. Math.* 45 (1914), 1175, a note of G. Armellini (*Comptes Rendus* 158 (1914), 253–255) is erroneous. Cf. also T. Levi-Civita, *Ann. di Mat.* (3) 9 (1903), 1–32.

§462–§476: Some of these results may be thought of as refinements for the limiting case of the restricted problem of the corresponding facts concerning the planar problem of three bodies. Cf., in particular, §464 and §469, §474–§476 with §358–§359, §380–§382, respectively (as to references, cf. those given above in connection with §382). Correspondingly, it seems to be hard to give exact references to the literature of all the facts collected in §462–§467 bis, where the presentation is simpler and more complete than usually given; cf. M. H. Martin, *Amer. Journ. of Math.* 53 (1931), 167–174 and Natalie Rein, *ibid.*, 58 (1936), 735–736. The table of §468 was calculated by Jenny E. Rosenthal (*Astr. Nachr.* 224 (1931), 169–172, where the heads of the last two columns must obviously be interchanged). Hill's curves of zero velocity were transferred from his limiting case (§495–§497) to the case of the actual restricted problem (§471–§473) by K. Bohlin (*Bihang Stockh. Akad.* 13 (1887), no. 1; *Acta Math.* 10 (1887), 115–118, where Hill is not mentioned). A detailed study of these curves was given for $\mu = 1/11$ by Sir G. H. Darwin (1897; *Papers* 4, 6–12); while G. Kobb (*Bull. Astr.* 18 (1901), 219–221; 25 (1908), 411–415) gave applications to the case of minor planets, where the mass ratio is that corresponding to Jupiter and the Sun. The solutions of the linear equations (19) were considered in the case of characteristic exponents of stable type by C. V. L. Charlier (*Öfv. Stockh. Akad.* 57 (1900), 1059–1082; cf. the corrections of N. Moisseiev, *Revista Univ. San Marcos* (Lima), 1937, no. 421), in the case of the unstable equilateral type by E. Strömgren (*Astr. Nachr.* 168 (1905), 105–108). E. Strömgren has also studied (*Medd. Danske Akad.* 10 (1930), no. 11) the coalescence of these two types in the latter case. The appearance of secular terms in the limiting case (cf. the end of §476) seems to be the first example of such an occurrence in a linear conservative dynamical system and was pointed out by A. Wintner, *Math. Ztschr.* 32 (1930), 660–661.

For a detailed discussion of the solutions of (19) cf. also M. Martin, *Astr. Nachr.* 244 (1931), 161–170.

§477–§477 bis: Because of the extreme simplicity of the misleading case of §477 bis, the problem formulated in §477, which seems to be very difficult, is usually overlooked. The short, though quite special consideration in §477 bis (cf. L. Fejér, *Crelle's Journ.* 131 (1906), 216–233), which is independent of the general criterion of §133 and also of the classical theory of solutions asymptotic to a position of equilibrium (Poincaré, Liapounoff, Hadamard), is only an adaptation of the considerations of Jacobi (§332).

§478–§488: The elementary solutions of §479 were pointed out by Levi-Civita (cf. G. Pavanini, *Ann. di Mat.* (3) 13 (1906), 184–192). The question indicated in §483 bis goes back to Newton's *Principia* and has led, two centuries later, to Adams' introduction of infinite determinants (cf. §524 below). The presentation in §480–§482 follows that given by Levi-Civita (*Ann. Éc. Norm. Sup.* (3) 28 (1911), 325–376), who derived (ibid.) the result of §487 in a less sharp form, by proving that the remainder term of the linear approximation to the node is bounded. The almost periodicity of this remainder term (§485–§487) was then observed by Wintner, *Ann. di Mat.* (4) 10 (1932), 277–282; cf. *Amer. Journ. of Math.* 62 (1940), 49–60. The considerations of Levi-Civita were extended to the actual (instead of the restricted) problem of three bodies by Libera Trevisani (Mrs. Levi-Civita), *Atti. Ist. Veneto* 71₂ (1912), 1089–1137. Cf. also Emma Trapani, *Rend. Acc. Napoli* (3) 25 (1919), 48–69. The existence of a mean motion and the almost periodicity of the remainder term in the general theorem of §484 were formulated by Wintner as a conjecture and subsequently proved by Bohr (*Medd. Danske Akad.* 10 (1930), no. 10₁; cf. *Comm. Math. Helv.* 4 (1931), 51–64, where the preservation of the moduli is proved). The elegant remarks of §488 are due to Levi-Civita, loc. cit. 352–353; cf. also *Acad. Polyt. Ann. do Porto* 12 (1912), 193–206.

§489–§502: G. W. Hill's fundamental paper (*Works* 1, 284–335), dealing with the case (4), appeared in 1878. The curves of zero velocity (§495–§497) were introduced by Hill (loc. cit.) in order to prove that the distance between the Earth and the Moon must remain bounded from above for all time, if the motion is defined by (1₁) and (4). The characteristic exponents of the solutions of equilibrium (§494) were considered by Poincaré (*Méth. Nouv.* 1 (1892),

159–161). The regularization of Levi-Civita (1904; cf. §447–§450) was applied to Hill's limiting case (§498) by Birkhoff (Palermo Rend. 39 (1915), 314–315). The result of §501 bis was proved by Poincaré (Acta Math. 13 (1890), 74–79; cf. K. Bohlin, *ibid.* 10 (1887), 115–117) by a less straightforward consideration. Incidentally, while all this remains valid if (4) is replaced by (2), everything breaks down if there is more than one free particle, as in case of the actual problem of three bodies (cf. Bohlin, *loc. cit.*, 118–121; Poincaré, Méth. Nouv. 3 (1899), 165–174). This situation may be thought of as connected with questions of transitivity. The result derived in §500 from Levi-Civita's regularization was first obtained by Birkhoff (*loc. cit.* 284–285), for (2) instead of for (4), by using his own regularization (§453); correspondingly, he was able to determine (*loc. cit.* 285–287) the topological structure of the isoenergetic phase space also in the remaining three of the four general types described in §472. The applicability of the ergodic theorem, emphasized in §501 bis (cf. Wintner, Math. Ztschr. 36 (1933), 637), is due to the fact that the asymptotic distributions involved (§123–§124) happen to remain unaffected by isoenergetic phase *and* time transformations of the type considered in the footnote to §49.

It is interesting that, while Poincaré recognized the fundamental character of Hill's investigations immediately, the other leading contemporaneous authority in mathematical astronomy, Bruns, who reviewed Hill's work in the Fortschr. d. Math. (10 (1878), 782), does not appear to have been impressed.

§503–§515: The papers mentioned in the references to §305–§307 apply three different, though in principle equivalent, analytical methods for existence proofs of periodic solutions of simple type in case of a general dynamical system: (i) the method of analytic continuation, based on Cauchy's local existence theorems (Poincaré); (ii) the integral technique of successive approximations and Green functions (Lichtenstein); (iii) the method of comparison of undetermined Fourier coefficients, a method depending on existence theorems for non-linear infinite implicit systems of equations. The method of Hill (*loc. cit.*) is this third method (cf. §505–§506), although he emphasized (cf. *loc. cit.*, p. 287, the section: "I regret that on account of the difficulty of the subject . . . it does not appear that anything in the writings of Cauchy will help us to the conditions of convergence") that he was unable to give the necessary existence (or convergence) proof. Such an existence proof (§507–§515) was sub-

sequently given by Wintner, *Math. Ztschr.* 24 (1925), 259–265. It should be mentioned that, according to very elementary considerations of Birkhoff (*loc. cit.*, 316–317), the existence of the periodic solutions is a much easier question in the retrograde case $m < 0$ than in Hill's case $m > 0$. According to a short review in the *Fortschr. d. Math.* 26 (1895), 1103, a proof for the convergence of Hill's trigonometrical series was given by Liapounoff in a Russian paper (1895).

§516: As to the details of the complete induction concerning the m -factors of a_j/a_0 , cf. H. Poincaré, *Leçons de Méc. Cél.* 2₂ (1909), 35–36. The criterion of O. Hölder (*Sächs. Sitzber.* 63 (1911), 388–393) may be proved in the same way as its analogue (or, rather, generalization) for the case of Fourier-Stieltjes transforms, in which case the criterion has been applied often recently in proofs for the smoothness of certain distributions.

§517: Notwithstanding the nearly circular character of the orbit, the treatment of this principal lunar inequality (which is the “variation” in the nomenclature of lunar theory and was considered already in Newton's *Principia*) presented, until Hill's work, one of the principal obstacles to a satisfactory approach to the analytical description of the path of the Moon.

§518: Cf. Wintner, *Math. Ztschr.* 30 (1929), 211–227. The connection of the “Euler transformation” (§518 bis) with the older lunar theories is discussed by Hill, *loc. cit.*, 315–316.

§519: Originally, Hill (*loc. cit.*, 326) made a curiously incorrect statement concerning the continuation of his cuspidal orbit (*loc. cit.*, 328–335). Afterwards he mentioned in his *Coll. Works* (*loc. cit.*, p. 326), that the correct situation was pointed out to him, before Poincaré (*Méth. Nouv.* 1 (1892), 105–109), by Adams (apparently unpublished). A path in which the small loops resulting from the cusps became considerable was calculated in 1892 by Lord Kelvin (*Papers* 4, 520). Cf. also K. Matukuma, *Proc. Imp. Acad. Jap.* 6 (1930), 6–8, 9 (1933), 364–366 (and 8 (1932), 147–150, where the retrograde paths are considered).

§519 bis: For a detailed discussion of Strömgren's empirical principle, cf. Wintner, *Die Naturwiss.* 19 (1931), 1008–1017; *Bull. Astr.* (2) 9 (1936), 251–253. The way in which E. Strömgren arrived at this principle is discussed by him, e.g., in his report, *Bull. Astr.* (2) 9

(1933), 87–130, where detailed references are given to the numerical investigations at the Copenhagen Observatory. The mathematical proof for the truth of Strömgren's empirical principle was given by Wintner, *Math. Ztschr.* 34 (1931), 321–349. For a short presentation of essentially the same proof, cf. G. D. Birkhoff, *Pisa Ann.* (2) 5 (1936), 39–42. The validity of the termination principle may be followed explicitly in integrable cases (cf. P. Stäckel, *Math. Annalen* 42 (1893), 537–563, *passim*); while certain partial statements contained in Strömgren's universal formulation occur in the literature before him in certain non-integrable cases also (cf., in particular, Birkhoff, *Trans. Amer. Math. Soc.* 18 (1917), 257–258, where reference is made to Poincaré).

§520: In principle, though not in detail, all this dates back to Hill (1877; *Works* 1, 244–251); cf. H. Poincaré, *Bull. Astr.* 17 (1900), 87–104, A. Wintner, *Amer. Journ. of Math.* 53 (1931), 611–616.

§521–§522: A. Wintner, *Amer. Journ. of Math.* 59 (1937), 795–802.

§523–§524: G. W. Hill, *loc. cit.* 252–270; J. C. Adams, *Papers* 1, 181–188 (1877); 2, 85–103 (posthumous). The mathematical justification of the Adams-Hill method of infinite determinants is due to Poincaré (*Bull. Soc. Math. de France* 14 (1886), 77–90; *Méth. Nouv.* 2 (1893), 260–267, where Hadamard's theory of entire functions is used, and *Bull. Astr.* 17 (1900), 134–143, where a rather concise treatment is given; cf. also *Lecons de Méc. Cél.* 2₂ (1909), 44–57). For further references cf. the report of H. Burkhardt, *Int. Math. Congr. Chicago* (1893) *Papers*, 1896, pp. 13–34.

§525: The difficulties involved in a consistent application of the method of infinitely many variables are hardly different from the problem of “small divisors” in classical celestial mechanics; cf., in fact, Wintner, *Math. Annalen* 96 (1926), 303, and *Math. Ztschr.* 30 (1929), 214–215. As pointed out recently (Wintner, *Proc. Nat. Acad. Wash.* 26 (1940), 127), these classical difficulties of perturbation theory may be thought of as being identical with the modern problem of irrational rotation numbers (cf., in particular, Birkhoff, *Ann. Inst. Poincaré* 2 (1932), 369–386; *Bull. Amer. Math. Soc.* 38 (1932), 374–375). It is clear from Birkhoff's investigations that the problem is actually one concerning integrability (correspondingly, cf. a verification carried out in certain integrable cases by J. Horn (Crelle's

Journ. 126 (1903), 194–232), who also gave (ibid. 131 (1906), 224–245) a clear arrangement of the formal calculations involved in the corresponding non-integrable case, both times in the neighborhood of a solution of equilibrium). The older literature of formal trigonometric expansion in celestial mechanics is collected on pp. 61–79 of Marcolongo's bibliography. A modern treatment of these formal expansions is due to Birkhoff (Amer. Journ. of Math. 49 (1927), 1–38 and Dynamical Systems (1927), Chap. IV; cf. also ibid., Chap. III, and Acta Math. 43 (1920), 1–79).

§526–§529: The reduction of §526 to §529 by means of the theory of almost periodic functions is that given by A. Wintner, Math. Ztschr. 31 (1929), 434–440. The papers referred to in the footnotes to §529 are H. Bruns, Astr. Nachr. 109 (1884), 215–222 [concerning Borel series (1894) and Baire categories (1899), cf., e.g., H. Hahn, Theorie der reellen Funktionen (1921), 313–317 and 75–82, 99–109; for an instance of the Baire argument before Bruns, cf. H. Hankel (1870), no. 153 (1905) of Ostwald's Klass., pp. 95–98] and H. Gylden, Comptes Rendus 106 (1888), 1584–1587, Öfv. Stockh. Akad. 45 (1888), 77–87, 349–358. In connection with these footnotes, it is interesting to observe that, in view of the investigations of T. Brodén (cf. Öfv. Stockh. Akad. 57 (1900), 239–266 and his paper discussed by H. Hahn, loc. cit., 311–313), also the celebrated measure principle of “either 0 or 1” probability in the theory of real functions (cf., e.g., P. Hartman and R. Kershner, Amer. Journ. of Math. 59 (1937), 809–822) has Gylden's considerations as its starting point; so that even the theory of measure on an infinite product space may be considered as having an astronomical origin. For further references to the literature of small divisors, cf. A. Wintner, loc. cit.

For a short mathematical introduction into the formal foundations of modern lunar theory, cf. H. Poincaré, Bull. Astr. 17 (1900), 167–204, where, however, the purely analytical point of view predominates. The astronomical point of view is less in the background in the lectures of J. C. Adams (Papers 2, 1–84) and Sir G. H. Darwin (Papers 5, 16–58), which, because of their concise clearness, can be recommended as introductions to the practical problems in lunar theory. The standard text-books of this theory are vol. 3 (1894) of Tisserand's Mécanique Céleste, E. W. Brown's Treatise (1896), and, from a less astronomical point of view, vol. 2₂ (1909) of Poincaré's Leçons.

INDEX

The numbers refer to the pages (pp. 3–410)

- Action integral 12, 71, 80, 124, 126;
 - isoenergetic, 72, 85, 124, 126, 132;
 - for periodic functions, 74, (169), 133, 137; for two bodies, 181, 183, 188
- adiabatic invariants 97
- almost periodic functions 140, 141, 377, 379, 405, 408
- angular momentum 148, 151, 178, 234, 315, 375; elimination of, 316, 319, 321, 345; relative (synodical), 222
- angular variables, cf. torus
- anomaly, eccentric, 195, 203; mean, 196, 203; true, 196, 203
- areal velocity 181; cf. angular momentum
- asymptotic values 218
- Barycentric, coordinates, 242, 350; chains, 309
- Bessel functions 204, 221
- biopolar coordinates 42, 144–145, 183, 351, 354–355; rational, 40, 355
- brackets 18–22
- broken extremals 122, 131, 174, 189, 190, 198
- Canonical, conjugates (mates), 14; extensions, 35; integration constants, 76, 158, 161; variation of, 78; matrices, 45, 110
- canonical transformations 22; criteria for, 23, 27, 31, 33, 34; remainder functions and multipliers of, 24; completely, 26; binary, 29; extended, 35
- central, configurations, 273, 279, 295; forces, 148, 151
- centre, of mass, 234; of force, 323; equation of, 197
- centrifugal forces 238, 249, 286, 319, 321, 349, 351
- characteristic exponents 105; in the dynamical case, 109; for constant coefficients, 67, 107, 110, 136
- characteristic partial diff. equ. 78; isoenergetic, 82; for one degree of freedom, 131; for a central force, 157; 183
- class, $C^{(n)}$, 4; $C^{[n]}$, 5
- classical integrals 20–21, 97, 240, 241; cf. 351
- collinear, solutions, 249, 317, 344; central configurations, 273, 276, 278; homographic, 249, 287, 299; homothetic, 287, 299; of relative equilibrium, 303, 366, 371, 382; characteristic exponents of, 305, 371, 382
- collisions, binary, 198–199, 265, 267, 271, 328–334, 353, 385; simultaneous (or general), 253, 279, 334–338; rectilinear, 248; continuable and non-continuable, 339–342; cf. real singularities
- compactness, 58, 63, 326, 333, 339–340, 356–359, 385
- complete solutions 79; isoenergetic, 83
- complex singularities 213, 215, 217; 221; 325
- configurations 273
- configuration space 14
- conformal transformations 36–39, 164; 147, 149, 163
- conjugate points 126, 180, 188–192
- conservation principles, cf. classical integrals
- conservative 17
- constant of gravitation 233
- contact transformations 8, 31
- contragradience 48; cf. brackets
- Copenhagen, numerical work of Observatory, 350, 355, 403
- Coriolis forces 113, 150, 165, 222, 238, 249, 286, 319, 321, 349, 351, 372
- critical commensurabilities 226
- critical points, cf. equilibrium, index relations
- cubic law of attraction, 115, 117, 200, 290, 292, 294
- curl 4

- curvature, of a curve, 167, 179; of a surface, 151, 163, 165 (in the problem of two bodies, 180)
- cusps 62, 119, 122, 124, 131, 174, 180, 228, 354, 403
- cyclic, *cf.* ignorable
- Degree of freedom 14; one 131; two 161; *cf.* reduced
- Diophantine approximations 93, 96, 98, 140, 146, 153, 224, 228, 378, 405, 407–410
- direct orbits 152, 179, 223, 224, 226, 232, 401
- displacements (infinitesimal) 65; isoenergetic 75, 170; normal 171
- distribution function 89; asymptotic, 90
- divergence 59
- domain 4
- dynamical similarity 115, 116, 235; *cf.* Kepler's third law
- dynamical systems 112
- Elliptic coordinates, *cf.* bipolar coordinates
- energy 9, 14, 68, 71, 112, 151, 152; kinetic, 113; potential, 113; relative (synodical), 223, 351
- equation of centre 197
- equations of variation, *cf.* Jacobi equations
- equilateral, configurations, 277, 324; solutions, 300; of relative equilibrium, 304, 345, 366; characteristic exponents of, 305, 371
- equilibrium solutions (points) 62, 119–121, 124; Jacobi equations of, 66, 101; stability of, 99–101, 372–373; characteristic exponents of, 67, 107, 110; for circular paths, 154–156; in the restricted problem, 366, 371; in the lunar case, 382; *cf.* relative equilibrium
- ergodic, theorem 90, 387; recurrence, 91
- Eulerian, angles, 56, 158–161, 320 375; centra, 145, 351; top, 145; transformation, 402
- Flat, central configurations, 273, 278; solutions, 245, 315; homographic 287
- flows 88; incompressible, 88; isoenergetic, 162, 167, 346, 387
- force function 113
- function groups 20, 22
- fundamental matrix 103
- Galilei group 240
- general solution 62; of linear systems, 102; of Hamiltonian systems, 77, 80; *cf.* flows
- geodesics 126, 127, 144, 147, 149, 151, 165, 169, 180, 182, 321
- gradient 4
- Halley's equation 200
- Hamiltonian, function 14; system (equations), 67
- heliocentric, coordinates, 257, 307; linear momenta, 308
- Hessian (determinant) 6; matrix, 4; polar, 33
- homographic solutions 284, 295, 298
- homothetic solutions 286, 299
- Ignorable coordinates (or momenta) 129, 142, 148, 152, 157, 259, 310, 317–320
- imprimitivity 97, 156
- inclination 56, 159, 161, 318, 343, 374
- incompressibility 26, 36, 88, 167
- independence, of functions, 17; of integrals, 62
- independent integrals 12, 73, (168), 183
- index relations 94, 169–170, 367
- inertial, coordinates, 233; transformations, 237
- "integrable" systems 144, 161–162, 349
- integrals 61–62; conservative, 62; isolating (uniform), 96, 155; conservation, *cf.* classical
- integral invariant (linear) 12, 73
- intrinsic equations 168, 343
- invariable plane 244, 247, 249, 253, 265, 271, 286, 312, 318, 329, 342
- invariant relations, sets, systems, 60; unrestricted, 85–86
- involution, functions in, 19, 27; integrals in, 68
- involutory transformations 7, 36
- irreversibility 113, 118, 130, 150, 220
- isosceles, configurations, 324; solutions, 262–265, 313–315, 344, 374

- Jacobian (determinant) 6; matrix, 4; matrix along a path, 64; constant and integral, 223, 350, 352, 382; identity, 18; last multiplier, 88
- Jacobi equations (systems of) 65; transformation rule of, 66; integrals of, 66, 75; in the Lagrangian and Hamiltonian cases, 74; of an equilibrium, 67; for a circular path, 154; for two degrees of freedom, 170-171
- Kapteyn's inequality 218
- Kepler's equation 197, 212-221
- Kepler's third law 179, 204, 380, 401
- kinematics, in a Euclidean space, 52; of continua, 16, 44
- Lagrangian, automorphisms, 11; derivatives, 9; their covariance, 10; equations, 69; their invariance, 70; function, 14; identity, 13; series, 215-222; system, 69; top, 145
- Lambert's, angle, 201; dictum, 143; theorem, 182
- Legendre transformation 6-8
- linear Hamiltonian systems 108-111
- linear momentum 240; elimination of, 258, 311
- Liouville's, condition, 88; systems, 138; theorem, 39
- loops 169, 176, 228, 229, 403
- Mass 233
- matrices 3-4, 43-44
- Maupertius' principle 124
- mean motion 377; 84, 135, 139, 153, 154, 203, 226, 379, 405
- minimal paths 92
- momenta, canonical, 14; cf. angular, linear, polar inertia
- monodromy, matrix, 104; group and its invariants, 105
- multiplier, of a canonical transformation, 24, 26; of a monodromy group, 105; Jacobi's last, 88-89
- multi-valued logics 144
- Newtonian attraction 154-157, 178, 233
- node 56, 159-161, 312, 375; elimination of, 320; line of, 312
- Orthogonal matrices, cf. rotations
- osculating coordinates 309
- Parabolic coordinates 40, 43, 193, 352, 385
- parentheses, cf. brackets
- particular solutions, cf. solutions
- pendulum 137
- periastron, cf. perihelion
- perigee 406
- perihelion 158, 161, 195; rotating, 153, 224-225
- period and energy 73, 116, 133, 135, 139, 146, 154-157, (168-169), 179, 226, 235, 302, 400
- Pfaffians 13, 31, 33, 71, 316, 320
- phase space 14
- planar solutions 245, 321
- polar coordinates 30, 37; cf. rotations
- polar factorization of a matrix 44, 46
- polar inertia momentum 234
- positive definite 44
- primitivity, cf. imprimitivity
- product, scalar, 4, 49; vector, 49
- product space 9
- parallax 381
- Poisson's parentheses, cf. brackets
- problem, of two centra, 145, 351; of two bodies 178, 261, 277, 296, 301; of several bodies, 233, 242; of three bodies, 346; restricted, 347 (non-planar, 373); lunar 381 (non-planar, 388)
- Radial symmetry 148-151
- real singularities 198-200; 326, 329, 338, 339; 356-359; 385
- reciprocal radii 36, 330
- rectilinear solutions 248, 249, 317, (149, 152); homographic, 287, 299; homothetic, 299
- recurrence theorem 91
- reduced, degree of freedom, 316-321, 343 (265, 296); manifold of states of motion, 344-346, 385-388
- region 87
- regular analytic Fourier series 210, 400
- regularization, cf. collisions, real singularities
- relative coordinates 243-244, 318-323
- relative equilibrium 286, 300, 366, 382; Jacobi equations of, 301, 370; characteristic exponents of, 305, 371, 382

- remainder function 24
- retrograde, cf. direct
- reversibility 113, 114, 127, 131, 138, 165, 169; 235, 321, 349
- rotations 44, 50–57

- Saturn, rings of, 306
- scalar 3
- secular perturbations, linear, 49, 99; non-linear, 99
- secular terms 67, 106, 108, 371, 404; cf. mean motion
- sidereal 222, 223, 224, 232
- singularities, cf. real, complex
- small vibrations 48, 67, 101–102, 110–111, 135–136
- solution (path), 59, 85
- stability, classical, 98; distributional, 90; cf. equilibrium
- stable type, characteristic exponents of, 67, 101, 105
- steepest descent, method of, 210–211
- Strömgren's principle 403
- surface transformations 95, 98, 162, 349
- synodical 222, 223, 224, 232, 350, 401
- syzygies 248, 324, 344; axis of, 350; field of force along, 359–366

- Tauberian theorems 255, 279
- time variable, absolute, 233; change of, 36, 115, 125, 127–128, 134, 138, 164, 193, 198, 203, 281, 322, 329, 331, 334, 343, 352
- trace 102
- torus 87, 95, 96, 98, 136, 140, 153, 162, 379
- transitivity, metrical and regional, 92, 95; cf. primitivity
- transversal 122, 173

- Uniformization (real) 134, 140, 146, 194, 197, 342; of collisions, 198–199, 329, 334, 342, 354, 385
- unrestricted, solutions (paths), 85, 341, 359, 385; invariant sets, 86

- Variation of canonical integration constants 78
- vectors 3; Euclidean, 49
- virial 73, 114, 235, 250, 252, 373

- Zero velocity, sets (curves) of, 120, 131, 166, 173, 177, 179, 180, 230, 232, 369, 384

